National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file - Votre référence

Our file - Notre référence

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

UNIVERSITY OF ALBERTA

CONFORMATIONAL STUDIES OF COLLAGEN TELOPEPTIDES AND BRADYKININ ANALOGS BY NMR AND CD SPECTROSCOPY

BY

XIAOHONG LIU

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

DEPARTMENT OF CHEMISTRY

EDMONTON, ALBERTA

FALL, 1993

National Library of Canada
Acquisitions and Bibliographic Services Branch
395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale du Canada
Direction des acquisitions et des services bibliographiques
395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-88043-0

Canada

University of Alberta
Edmonton
Department of Chemistry
Faculty of Science
23-41 Chemistry Bldg. Tel: (403) 492-5231 Fax: (403) 492-8231

To: Faculty of Graduate Studies
University of Alberta

I give permission to use the paper entitled "The Aggregation Properties of Some Bradykinin Analogs" which is in press in the Journal of Biomolecular Structure and Dynamics for use in the Ph.D. thesis by Xiaohong Liu, June (1993) in the Department of Chemistry at the University of Alberta.

John M. Stewart John R. Cann Lajos Gera

George Kotovych

June 1, 1993

GK/co

University of Alberta
Edmonton

Department of Chemistry
Faculty of Science

To: Faculty of Graduate Studies
University of Alberta

I give permission to use the paper entitled "Proton Magnetic Resonance Studies of Bradykinin Antagonists" which is in press in *Biopolymers* for use in the Ph.D. thesis by Xiaohong Liu, June (1993) in the Department of Chemistry at the University of Alberta.

Xiaohong Liu John M. Stewart Lajos Gera George Kotovych

June 1, 1993

GK/co

University of Alberta
Edmonton
Canada T6G 2G2

Department of Chemistry
Faculty of Science
E3-44 Chemistry Bldg., Tel. (403) 492-3254 Fax (403) 492-8231

To: Faculty of Graduate Studies
University of Alberta

I give permission to use the paper entitled "A sequence-Dependent ^{1}H-NMR Study on the Formation of β-Turns in Tetrapeptides Containing Charged Residues" published in *Biopolymers* **32**, 119-130 (1992) for use in the Ph.D. thesis by Xiaohong Liu, June (1993) from the Department of Chemistry at the University of Alberta.

Xiaohong Liu | Paul G. Scott | Albin Otter | George Kotovych

June 1, 1993

GK/co

University of Alberta
Edmonton
Canada T6G 2G2

Department of Chemistry
Faculty of Science
E3-44 Chemistry Bldg., Tel. (403) 492-3234 Fax (403) 492-8231

To: Faculty of Graduate Studies
University of Alberta

I give permission to use the paper entitled "Solution Conformation of the Type I Collagen Alpha-2 Chain Telopeptides Studied by ^{1}H and ^{13}C NMR Spectroscopy" published in the *Journal of Biomolecular Structure and Dynamics* **8**, 63-80 (1990) for use in the Ph.D. thesis by Xiaohong Liu, June (1993) in the Department of Chemistry at the University of Alberta

Xiaohong Liu | Paul G. Scott | Albin Otter | George Kotovych

June 1, 1993

GK/co

UNIVERSITY OF ALBERTA
RELEASE FORM

NAME OF AUTHOR:	Xiaohong Liu
TITLE OF THESIS:	Conformational Studies of Collagen Telopeptides and Bradykinin Analogs by NMR and CD Spectroscopy
DEGREE:	Doctor of Philosophy
YEAR THIS DEGREE GRANTED:	1993

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly, or scientific research purposes only.

The author reserves other publication rights and other rights in association with the copyright in the thesis, and except as hereinbefore provided neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form without the author's prior written permission.

Xiaohong Liu

(Student's Signature)

(Student's Permanent Address)

Department of Chemistry
University of Alberta
Edmonton, Alberta
Canada T6G 2G2

Date: Aug 9, 1993

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled "CONFORMATIONAL STUDIES OF COLLAGEN TELOPEPTIDES AND BRADYKININ ANALOGS BY NMR AND CD SPECTROSCOPY" submitted by XIAOHONG LIU in partial fulfillment of the requirements for the degree of DOCTOR OF PHILOSOPHY.

Dr. George Kotovych

Dr. R. E. D. McClung

Dr. Loren Hepler

Dr. John C. Vederas

Dr. Paul G. Scott

Dr. Hans Vogel

Date: Aug 2, 1993

ABSTRACT

The N- and C-terminal telopeptides of the α-2 chain of type I collagen in CD_3OH/H_2O solutions have been studied using high-field ^1H and ^{13}C NMR. The conformation of the N-telopeptide (nonamer) is predominantly extended with a small proportion of the molecules existing in a type I β turn. The four residues involved in this turn are D^3-A^4-K^5-G^6 which is stabilized by a C = O(D^3) - NH(G^6) hydrogen bond. The C-terminal telopeptide is extended throughout. A model is proposed involving charge-charge and hydrophobic interactions between the extended α-2 chain N-telopeptide and the adjacent segments of the triple-helix. A similar model is proposed for the C-telopeptide.

The conformations of the type II and type III collagen α-1 chain C-telopeptides in CD_3OH/H_2O (80/20) solution have been studied by means of two-dimensional proton NMR and circular dichroism spectroscopy. The conformation of the type II C-telopeptide is mostly extended. Evidence from CD spectroscopy suggests that a minor proportion of the peptide might be helical (ca 8%). The conformation of the type III C-telopeptide is mostly extended except for a β-turn ranging from Gly^8 to Glu^{11}, which is stabilized by a hydrogen-bond between the NH group of Glu^{11} and the carbonyl group of Gly^8.

The importance of side-chain charge interactions in the formation of β-turns has been examined. Sixteen protected NAc-tetrapeptide amides were studied. The results indicate that a small proportion of type I β-turn exists in solutions of DEKS and DERS in methanol/water (60/40), while NEKS has an even smaller population of this turn. The other tetrapeptides are present in solution only in the extended conformation. These results clearly show the importance of the salt-bridge between the side-chains of K^2 and E^3 or R^2 and E^3 as well as the importance of the charge on the side-chain of the first residue

in stabilizing the β-turn. The relevance of statistical predictions for β-turns in short peptides is also discussed.

The solution conformations of DArg^0-[Hyp^3, Thi^5, DCpg^7, Cpg^8]-BK and DArg^0-[Hyp^3, DCpg^7, Cpg^8]-BK, two extremely potent bradykinin antagonists and two related antagonists which differ from each other only by one or two residues, were investigated by ^{1}H NMR in a methanol/water (80/20 v/v) solution. A turn-like structure, which coexists with the extended conformation, was observed between residues 2 and 5 for the most active antagonists, in direct correlation with the peptide activities. The data suggest that a turn-like structure between residues 2 and 5 could be important for antagonist activity and that the conformational difference between bradykinin agonists and antagonists is due to a turn formed at *different positions* in the peptides.

Finally, the aggregation properties for two bradykinin analogs, namely DArg-[Hyp^3, Thi^5, DSer^6, DCpg^7, Cpg^8]-BK [I] and DArg-[Hyp^3, DSer^6, DCpg^7, Cpg^8]-BK [II] were investigated by ^{1}H NMR and CD[75]. These two molecules have exhibited an abnormal, non-linear temperature dependence for the amide NH proton of Cpg^8. The NH of Arg^9 also shows a slightly non-linear temperature dependence at temperatures above 25 °C. In addition, a very slow exchange rate for the NH protons of DCpg^7, Cpg^8 and Arg^9 indicated aggregation of these two analogs which was confirmed using circular dichroism.

ACKNOWLEDGEMENTS

I would like to acknowledge some of the many individuals who have helped me with this research over the past six years. Firstly, I would like to thank my supervisor, Dr. George Kotovych, who has given me the opportunity to carry out research in his laboratory, and who has given me intellectual direction and emotional support over the years.

I would like to thank Dr. Albin Otter for his advice, help and warm friendship. I would also like to thank Dr. Paul Scott for his advice, his efforts in peptide synthesis and purification, and for his interest in collaborating on the projects. I am indebted to Dr. John Stewart and Dr. John Cann for their efforts in peptide synthesis and purification, for running CD spectra, and for their interest in collaborating on the projects. I would like to give my appreciation to Dr. Tom Nakashima and other members of the NMR lab for technical assistance.

Finally, I would like to thank my wife and my parents, for their gentle love, support and sacrifice over the years.

This research has been supported by the Alberta Heritage Foundation for Medical Research in the form of a studentship. Financial support by the Natural Sciences and Engineering Research Council of Canada, the Canadian Medical Research Council, and the Department of Chemistry, University of Alberta is gratefully acknowledged.

TABLE OF CONTENTS

Chapter	Page
1. Introduction	**1**
Collagens and Telopeptides	1
Bradykinin and its Analogs	6
Secondary Structures in Small Peptides	10
Spectroscopic Methods to Detect Secondary Structures in Small Peptides	15
References	30
2. Solution Conformation of the Type I Collagen Alpha-2 Chain Telopeptides Studied by ^{1}H and ^{13}C NMR Spectroscopy	**46**
Introduction	46
Materials and Methods	47
Results	50
Discussion	53
References	57
3 Conformational Analysis of the Type II and Type III Collagen α-1 Chain C-Telopeptides by ^{1}H NMR and Circular Dichroism Spectroscopy	**79**
Introduction	79
Materials and Methods	80
Results	85
Discussion	88
References	92

4. A Sequence-Dependent ^1H NMR Study on the Formation of β-Turns in Tetrapeptides Containing Charged Residues 107

Introduction 107

Materials and Methods 108

Results 113

Discussion 116

References 121

5 Proton Magnetic Resonance Studies of Bradykinin Antagonists 144

Introduction 144

Materials and Methods 146

Results 148

Discussion 153

References 157

6 The Aggregation Properties of Some Bradykinin Analogs 170

Introduction 170

Materials and Methods 171

Results and Discussion 174

References 180

7. General Discussion and Conclusions 192

References 201

LIST OF TABLES

Table		Page
Table 1.1	Dihedral Angles of Hydrogen Bonded β and γ Turns	44
Table 1.2	Characteristic Short Distances (Å) and Coupling Constants (Hz) in Some Regular Secondary Structures	45
Table 2.1	Summary of Experimental Parameters Used in the Two-Dimensional NMR Experiments	60
Table 2.2	Proton Chemical Shifts, Coupling Constants and NH Temperature Dependence of pE-F-D-A-K-G-G-G-P-NH_2.	62
Table 2.3	^{13}C Chemical Shifts of pE-F-D-A-K-G-G-G-P-NH_2 as a Function of pH.	64
Table 2.4	Proton Chemical Shifts, Coupling Constants and NH Temperature Dependence of NAc-S-G-G-Y-E-F.	66
Table 2.5	^{13}C Chemical Shifts of NAc-S-G-G-Y-E-F.	67
Table 3.1	^{1}H Chemical Shifts, Coupling Constants and NH Temperature Coefficients of the Type II Collagen α-1 Chain C-Telopeptide	95
Table 3.2	^{1}H Chemical Shifts, Coupling Constants and NH Temperature Coefficients of the Type III Collagen α-1 Chain C-Telopeptide	97
Table 4.1	Summary of Experimental Parameters Used in the Two-Dimensional NMR Experiments	124
Table 4.2	Proton Chemical Shifts, Coupling Constants and NH Temperature Coefficients of the Peptides Studied	126

Table 4.3 Summary of ^1H Chemical Shifts, Coupling Constants and NH Temperature Coefficients of DEKS, DERS and NEKS in CD_3OH/H_2O (60/40) 138
Table 4.4 Probabilities of β-turn Formation in Tetrapeptides 141
Table 5.1 Peptides Studied 160
Table 5.2. Pharmacological Data and NH Temperature Dependence Coefficients at Positions 5 and 9 for Peptides I-IV. 160
Table 5.3 Proton Chemical Shifts, Coupling Constants and NH Temperature Coefficients of the Peptides Studied 161
Table 6.1 Peptides Studied 183
Table 6.2 Proton Chemical Shifts, Coupling Constants and NH Temperature Coefficients of the Peptides Studied 184
Table 7.1 A List of New Bradykinin Antagonists 203
Table 7.2. Pharmacological Data for Peptides I-IV listed in Table 7.1. 203

LIST OF FIGURES

Figure		Page
Figure 1.1	Schematic representations of a Type II collagen monomer structure and telopeptide sequences in Type I, Type II and Type III collagens.	40
Figure 1.2	Standard nomenclature for the atoms and the torsion angles along a polypeptide chain.	42
Figure 1.3	Pulse sequences for some 2D NMR experiments.	43
Figure 2.1	The structures of the nonamer and the hexamer, namely the N-terminal and C-terminal Type 1, α-2 chain telopeptides.	68
Figure 2.2	The pH titration of the nonamer in aqueous solution.	69
Figure 2.3	The COSY spectrum of the nonamer.	70
Figure 2.4	The ROESY spectrum of the nonamer.	71
Figure 2.5	The low-field ^{13}C spectrum of the nonamer as a function of pH.	73
Figure 2.6	The ^{1}H-^{13}C correlation spectrum of the nonamer.	75
Figure 2.7	The results of the proton NMR data in CD_3OH/water, including the ROE data, of the nonamer, together with the two proposed structures.	77
Figure 2.8	Alignment of N- (A) and C-terminal (B) α-2 chain telopeptides with sequences of triple-helices displaced axially by 1D (234 residues) and 4D (936 residues)	78
Figure 3.1	Circular Dichroism spectra of the type II collagen α-1 chain C-telopeptide at 10 °C	99

Figure 3.2 The contour plot of a phase-sensitive 500 MHz TOCSY experiment of the type II collagen α-1 chain C-telopeptide at a mixing time of 80 ms 101

Figure 3.3 Plot of several build-up curves of the pairs of protons in the type II collagen α-1 C-telopeptide indicated in the plot 102

Figure 3.4 Two sections of a 500 MHz phase-sensitive NOESY experiment of the type III collagen α-1 chain C-telopeptide at a mixing time of 200 ms 104

Figure 3.5 The amino acid sequences of the type II (top) and type III (bottom) collagen α-1 chain C-telopeptides with unique residues underlined. The observed sequential and medium range NOEs are indicated schematically by solid bars 106

Figure 4.1 NH-NH region of the 500 MHz phase-sensitive ROESY experiments on DEKS, DERS and NEKS in CD_3OH/H_2O (60/40) 143

Figure 5.1 The primary structure of peptide I 166

Figure 5.2 The contour plot of a phase-sensitive 500 MHz TOCSY experiment of peptide I in CD_3OH/H_2O(80/20) 167

Figure 5.3 The NH/NH region of a 500 MHz phase-sensitive NOESY experiment of peptide I at a mixing time of 300 ms 168

Figure 5.4 The amino acid sequences and the NOE summary of peptides I and II 169

Figure 6.1 Plot of the chemical shifts of the amide protons of peptide II in CD_3OH/H_2O(80/20) vs temperature 188

Figure 6.2 Plot of part of the NH region of peptide I in CD_3OD/D_2O(80/20) vs time..... 189

Figure 6.3 Circular dichroism spectrum of peptide I in different concentrations in methanol/water (80/20) 191

LIST OF ABBREVIATIONS AND SYMBOLS

CD	circular dichroism
COSY	two-dimensional correlation spectroscopy
FID	free induction decay
HPLC	high performance liquid chromatography
LHCPL	left-handed circularly polarized light
mM	milli-molar
nm	nanometer
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser enhancement
NOESY	two-dimensional nuclear Overhauser enhancement spectroscoy
ORD	optical rotatory dispersion
ppb	parts per billion
ppm	parts per million
rf	radiofrequency
RHCPL	right-handed circularly polarized light
ROESY	two-dimensional rotating-frame NOESY
SDS	sodium dodecyl sulfate
TOCSY	two-dimensional total correlation spectroscopy

Naturally Occurring Amino Acids

A, Ala	Alanine
C, Cys	Cysteine
D, Asp	Aspartate
E, Glu	Glutamate
F, Phe	Phenylalanine

G, Gly	Glycine
H, His	Histidine
I, Ile	Isoleucine
K, Lys	Lysine
L, Leu	Leucine
M, Met	Methionine
N, Asn	Asparagine
P, Pro	Proline
Q, Gln	Glutamine
R, Arg	Arginine
S, Ser	Serine
T, Thr	Threonine
V, Val	Valine
W, Trp	Tryptophan
Y, Tyr	Tyrosine

Modified Amino Acids

Cpg	α-cyclopentaneglycine
Dhq	dehydroquinuclidine 3-carboxylic acid
Hyp	*trans*-4-hydroxy-L-proline
Igl	α-(2-indane)-glycine
Oic	octahydroindole-2-carboxylic acid
Thi	β-(2-thienyl)-L-alanine

CHAPTER 1

Introduction

Collagens and Telopeptides

Collagens are a family of chemically and structurally related, but genetically distinct, proteins that are found in the connective tissues of all multicellular organisms. The association of collagen fibers into bundles determines the architectural organization of the connective tissues and conditions their mechanical properties[1,2]. Of the fourteen or more collagen types that are found in higher vertebrates, type I, type II and type III, i.e., the fibril-forming interstitial collagens, are the most common ones and are axially aligned in staggered arrays to produce the bundles (fibrils) that comprise a major constituent of many structurally important biological materials, such as tendon, skin, and bone.

Many physical and physiochemical studies, including direct visualization of individual molecules in the electron microscope, have demonstrated the rod-like nature of collagen molecules. Molecules of collagen have a length of slightly less than 300 nm and a diameter of about 1.4 nm. The rod is neither rigid nor randomly flexible but appears to possess an intermediate level of flexibility which probably varies along its length. Collagen monomer is characterized structurally by a triple-helix which consists of three polypeptide chains (α chains), each a little over 1000 amino acid residues long. A prerequisite for the formation of this triple helix is a Gly-X-Y repeating tripeptide unit in the amino acid sequence of the three chains, where X and Y can be any amino acids but are often the imino acids proline and hydroxyproline. This sequence, with glycine in every third position and with an unusual abundance of hydroxyproline, forms the basis for

the chemical identification of collagen. The details of the three dimensional structure of the triple helix, established by X-ray diffraction and using known bond lengths and bond angles, have been reviewed[3]. At each end of the α-chains are short sequences without the repeating G-X-Y unit, called "telopeptides", which do not adopt the triple-helical conformation. Telopeptides constitute less than 4% of the complete collagen molecule, but contain the only tyrosyl residues of the molecule and the lysine residues, one in each telopeptide, which provide the potential sites for the formation of covalent cross links[4]. The formation of these covalent cross-links is the final step in the production of the stable functional unit of collagen - the fibril[5]. Inhibition of cross-link formation is cited as part of the underlying pathology in several connective tissue diseases[4,6]. Conversely, there is evidence that increased cross-linking is responsible for detrimental changes in the mechanical properties of fibrils during aging[7]. The telopeptides are sites of the major antigenic determinants of the molecule; they are the only sites for attack by any proteinases other than triple helix-specific collagenase. An outline of a basic collagen monomer structure is shown in Figure 1.1a.

Type I is the most abundant form of collagen and found as relatively thick fibrils, approximately 20 to 200 nm or more in diameter, depending on the tissue and age[8,9]. When organized into fibres and fibre bundles, they give mechanical strength to tissues such as dermis, tendon, bone and dentine. The three α-chains are non-identical in type I collagen, where the monomer comprises two identical α1(I) chains and a nonidentical α2(I) chain. The amino acid sequences of the α1(I) and α2(I) chains are known[10,11]. Species differences occur, but a substantial measure of homology between species exists for type I collagen[12]. Of 1050 amino acids in the α1 chain of calf skin type I collagen, 1014 occur in the repeating Gly-X-Y triple helix. The N-terminal telopeptide has 16 residues

and the C-terminal telopeptide 25 residues. The α2(I) chain has 1029 amino acids and telopeptides are 9 and 6 residues long for N-terminal and C-terminal, respectively. (Figure 1.1b).

Type II has three identical α-chains. Whereas the sequences of type I collagen telopeptides are highly conserved between species, there is considerable variability between the telopeptides of different genetic types of collagen types. Type II collagen is found in adults only in hyaline cartilage, where it is the dominant form, and in vitreous humour. *In vivo*, it is present as thin fibrils which are critical components in the physical integrity of cartilage. The N- and C- telopeptides of human type II collagen are 19 and 27 residues long, respectively[13] (Figure 1.1c). Type III is the second most abundant form of collagen in soft tissues other than cartilage. It is prevalent in tissues where distensibility is a major requirement for function, such as blood vessel walls, lung parenchyma, intestine and uterine wall[14]. *In vitro*, type III collagen has been shown to form thinner fibrils than type I and to reduce the diameter of copolymeric fibrils containing predominantly type I[15]. *In vivo*, it appears that the fibrils in many tissues contain both types I and III, but at different ratios in different organs and tissues[1,16]. The type III collagen monomer consists of three identical α-1 chains. Sequence data are available for the bovine N-telopeptide, which is 14 residues long[17] and the human[18] and chick[19] C-telopeptides which are 25 and 23 residues long, respectively. (Figure 1.1d)

With the aim of understanding how the mechanical properties of connective tissues can be related to the underlying molecular structures of collagen, collagen fibrillogenesis, the process of fiber assembly, has been intensively studied in hopes of elucidating the mechanism of this important step in the biogenesis of connective tissues. Most *in vitro* investigations have centered around the observation that collagen molecules have the ability to self-

assemble from solution, when the pH, temperature and ionic strength are adjusted to physiological values[20-22]. The fibrils produced *in vitro* are indistinguishable from those found in tissues[23-25]. It is now generally accepted that all the basic information needed to form the fibril is clearly contained in the collagen amino acid sequence and in the intrinsic properties of collagen *per se*, although other factors and other components of the extracellular milieu do play an important role in the process *in vivo*[12].

In vitro studies of the self-assembly of collagen fibrils have supported the concept[23] of the initial formation of "nuclei" which appear to be end-overlap structures[26,27]. The formation of "nuclei", or initial aggregates, is followed by lengthwise growth of the microfibril and subsequent lateral assembly into fibrils[28,29]. One aspect of the self-assembly process which was recognized very early was that proteolytic degradation of the telopeptides of collagen, without modifying the triple helical domain, would lengthen the fibril initiation and growth phases, or completely abolish the ability of the collagen molecules to self-assemble[28,30-32]. The addition of free amino-terminal telopeptide, isolated from the α1(I) chain, to a solution of native collagen, specifically enhanced the rate of fibril formation[33]. From these results it has been deduced that the N-terminal telopeptide of α1(I) chain directs linear growth of type I collagen fibril[30] while the C-telopeptide promotes both linear and lateral growth[31,34].

The functions of the telopeptides presumably depend on specific interactions between the telopeptides and the triple helices of collagen monomers adjacent to them in the growing fibrils. Because such interactions are likely to be strongly conformation dependent, it is now evident that a detailed knowledge of the conformations and interaction behavior of the telopeptides is an essential step toward a better understanding of many aspects of the fibril assembly process[8]. Since the telopeptides are very small compared to the

triple-helix and probably lack a well-defined periodic structure (see below), X-ray diffraction is unable to furnish any high resolution structure data; and electron microscopy can only indicate that the telopeptides in the fibril are in a condensed or folded conformation[8]. However, several attempts have been made to predict secondary structures from the amino acid sequences of telopeptides, and subsequently models for interaction between telopeptides and triple helices have been proposed[35-39]. One common feature of these predictions is that a loop exists around the lysyl residue which is involved in the intermolecular cross-link, but in most cases the predictions even from the application of the same algorithm, have often varied between laboratories and physical evidence available has, until recently, been insufficiently detailed to support or refute the various models proposed.

In this laboratory, the conformations of the telopeptides in solution have been studied using modern techniques of high-field NMR spectroscopy to provide more physical evidence and a better understanding of fibril assembly of collagens. So far, investigations on all the telopeptides in type I, II and III collagens[40-46] have been completed.

In Chapter 2, the high-field ^{1}H and ^{13}C NMR studies of the N- and C-terminal telopeptides of the α-2 chain of type I collagen in CD_3OH/H_2O solutions are described.

Chapter 3 presents the conformational studies of the type II and type III collagen α-1 chain C-telopeptides in CD_3OH/H_2O (80/20) solution by means of two-dimensional proton NMR and circular dichroism spectroscopy.

Chapter 4 discusses the importance of side-chain charge interactions in the formation of β-turns. Sixteen protected NAc-tetrapeptide amides were studied, namely the variants of DEKS; NEKS, EEKS, DDKS, DQKS, NQKS, DERS, NERS, EERS, DDRS, NDRS, DQRS, and DKES. Three tetrapeptides,

NPDM, NSDM and NDDS, were also studied as they have a high probability of forming β-turns, based on statistical predictions.

Bradykinin and its Analogs

Bradykinin, Arg-Pro-Pro-Gly-Phe-Ser-Pro-Phe-Arg, which may be considered as a prototype for the kinins and is naturally present in human body fluids (blood and urine), is known to possess extraordinarily pharmacological activities such as vasodilation and algesic activities[47-49]. It is one of the most potent vasodilators and causes dilation of blood vessels in muscle, kidney, viscera and various glands and also in the heart and brain[50]. Various smooth muscle preparations are found to contract in response to bradykinin. It is also a powerful algesic agent and causes an intense burning pain when applied to the exposed base of a blister, and a throbbing, burning pain in the hand when injected into the brachial artery. Recently, it has been suggested that bradykinin may be associated with the symptoms of the common cold[51,52].

In order to investigate the role that bradykinin plays in a variety of pathophysiological processes, the preparation of effective competitive antagonists is essential for the study of the actions of molecules involved in biological communication as in the case of bradykinin. Bradykinin can produce effects on its receptor which makes accurate evaluation difficult, if not impossible. In addition, rapid destruction of agonists causes serious problems in the estimation of their presence in biological systems and of their potencies in assays. Properly designed antagonists can overcome these problems[53], and the design and synthesis of potent, stable and specific bradykinin antagonists has long been considered a desirable goal in medicinal chemistry. Compounds with these characteristics could be important in the treatment of such diverse disorders as septic shock, asthma and rhinitis[49,54]. Because of the fact that

neither X-ray crystallographic nor NMR data pertaining to the receptor or ligand-receptor complex is available, many analogs of bradykinin have been synthesized in the period since the announcement of the structure of bradykinin in 1960, in order to increase the potency, selectivity and lifetime of action of bradykinin analogs and to obtain insight about their bioactive conformation. In almost all of these syntheses, analogs of natural amino acids were used to impose specific conformational constraints[55,56].

The first report of bradykinin analogs able to antagonize the action of bradykinin in the standard kinin assay system came in 1985, with the description of [DPhe^7]-BK (BK = bradykinin, [DPhe^7]-BK = Arg-Pro-Pro-Gly-Phe-Ser-DPhe-Phe-Arg)[57]. Before this, only agonists were obtained. Structures and biological activities of several hundred agonists were summarized in the reviews of Schroder[58] and Stewart[59]. The critical change conferring bradykinin antagonist activity upon analogs was the replacement of the Pro at position 7 with a D-aromatic amino acid residue. Acting on the assumption that modifications previously found on agonists to enhance receptor affinity might increase the potency of [DPhe^7]-BK, a weak antagonist, Thi (Thi = β-(2-thienyl)-L-alanine) was placed in positions 5 and 8, Hyp (Hyp = *trans*-4-hydroxy-L-proline) was placed in positions 2 and 3, and basic residues or dipeptides were added to the amino end of the molecule[53]. These manipulations were successful in converting the weak partial antagonist [DPhe^7]-BK to a much more potent analog with antagonist activity on all the standard assay tissues[60-62]. Among these early antagonists, DArg-[Hyp^3, DPhe^7]-BK and DArg-[Hyp^3, Thi^{5,8}, DPhe^7]-BK were the best. Now DArg-[Hyp^3, DPhe^7]-BK is in clinical trials for use as an antibradykinin drug in several pathological states.

The discovery of the first bradykinin antagonist established the necessary and, at times, sufficient requirement that a D-aromatic amino acid residue

replace the Pro residue at sequence position 7[53,57]. Additionally, the combination of a D-aromatic residue at position 7 with an L-Pro residue replacing the L-Phe at position 8 was also shown to produce antagonists[63]. Recently, bradykinin analogs with a modified D-Phe residue at position 7 combined with Pro derivatives at position 8 were demonstrated to be the most potent and long acting antagonists[64]. So far, bradykinin analogs having bulky, beta-branched or cyclic D-aliphatic residues at position 7 with bulky, or cyclic L-aliphatic residues at position 8, such as DArg^0-[Hyp^3, Thi^5, DCpg^7, Cpg^8]-BK (Cpg = α-cyclopentyl-glycine, see Figure 5.1 for the primary structure), have yielded the best antagonists[65,66], as have bradykinin antagonist dimers[67]. Thus, both the "conventional wisdom" that a D-aromatic residue is required at position 7 of BK antagonists, and structure-activity relationships leading to models for receptor interactions based on aromatic side chains at position 7 and/or position 8, must be altered[65].

The different biological activities of bradykinin analogs are believed to be conformation dependent. In the absence of a known receptor structure and a very small binding constant between bradykinin analogs and their receptors, the large number of bradykinin analogs, together with their structural and biological activities, provides a solid foundation for gaining insight into a possible bioactive conformation in solution. In various studies including ^{1}H NMR in aqueous solutions, bradykinin was found to exist in many conformational non-regular states[68]. However, bradykinin in less polar solvents such as dioxane[69] and SDS (sodium dodecyl sulfate) micelles[70] was found to possess a well-defined conformation characterized by a β-turn-like structure for residues 6 to 9. Earlier conformational studies in solution[68,69] on bradykinin itself and an antagonist also led to the hypothesis that the difference between an agonist and an antagonist is related to the type of β-turn adopted at the C-terminus, together with the required

orientation of the side chains[70]. To challenge this hypothesis, Kyle et al.[68,71] designed and prepared several constrained bradykinin analogs, which were assumed to have an inherently stabilized β-turn geometry at their C-terminus based on conformational analysis using empirical calculations. Two of them with big bulky side-chains at position 7 or 8, such as DArg^0-Arg^1-Pro^2-Hyp^3-Gly^4-Thi^5-Ser^6-DTic^7-Tic^8-Arg^9 (Tic = tetrahydroisoquinoline carboxylic acid residue), were found to be competitive antagonists, but NMR data were not reported for them. More recently, two bradykinin analogs containing α-MePro at position 3 or position 7 were reported to show reversed-turn conformations at both Pro^2-Phe^5 and Ser^6-Arg^9, respectively, in aqueous solution by NMR, but no data on activity were reported for these molecules[72].

In chapter 5, the solution conformations of [DArg^0, Hyp^3, Thi^5, DCpg^7, Cpg^8]-BK and [DArg^0, Hyp^3, DCpg^7, Cpg^8]-BK, two extremely potent bradykinin antagonists and two related antagonists which differ from each other only by one or two residues, were investigated by ^{1}H NMR, in order to provide more experimental evidence for the hypothesis that a β-turn in the four C-terminal amino acid residues of bradykinin analogs might be a prerequisite for high potency[68]. This study was carried out on the antagonists in the free state, as it is impossible to carry out NMR studies when they are bound to the receptor. These antagonists function at nanomolar concentrations, far below the limit of NMR detectability. The solvent for this study is a methanol/water (80/20 v/v) mixture, which is a structure-inducing solvent that has been used in the study of other small peptides[46].

In chapter 6, the ^{1}H NMR and CD studies of the aggregation properties for two bradykinin analogs, namely DArg-[Hyp^3, Thi^5, DSer^6, DCpg^7, Cpg^8]-BK [I] and DArg-[Hyp^3, DSer^6, DCpg^7, Cpg^8]-BK [II] (Cpg = α-cyclopentyl-glycine; Hyp = 4-hydroxy-L-proline, Thi = β-(2-thienyl)-L-alanine), were described.

Secondary Structures in Small Peptides

A certain conformation of a peptide can be described by a set of dihedral angles as indicated in Figure 1.2, since all of the bond lengths and bond angles are invariant, except perhaps for τ, the backbone N-C^{α}-C' angle, which is usually 110°, but could stretch to larger values in order to accommodate other strains in the structure[73]. The dihedral angle ω at the peptide is very close to 180° and the $C^{\alpha}(i)$ and $C^{\alpha}(i+1)$ are all *trans* to each other, because the peptide bond has partial (40%) double-bonded character. *Cis* peptide bonds, with ω=0°, can occur perhaps 25% of the time before proline residues but essentially rarely before any other residues. The proline ring is not quite flat, and occasionally protein structures are not being refined accurately enough to determine the direction of ring pucker. Considering the fact that dihedral angles for side chains are more difficult to define because of the greater flexibility of side chains, dihedral angles, ϕ and ψ are the source of essentially all the interesting variability in peptide conformation. Actually, secondary structures are usually refered to as regular arrangements of the backbone of the polypeptide chain without reference to the side chain types or conformations[74,75].

The value of each particular dihedral angle adopted by a residue in a molecule depends upon the overall energy of the molecule, which depends on bond angles and length, on the interactions between atoms not linked by covalent bonds and on geometrical restraints[74-76]. The permitted values of ϕ and ψ are generally indicated on a two-dimensional plot of ϕ-ψ angles known as the Ramachandran plot[77]. In a typical Ramachandran plot for a nonglycine, only 8% to 20% of the area in the plot is sterically allowed[73-75].

Many peptides have been found to adopt a few regular secondary structures that are also found in natural proteins. Although these regular

secondary structures in peptides and in proteins are homologous structures, the difference in molecular sizes leads to some distinct geometric and energetic characteristics for the regular secondary structures in peptides[55]. First, nearly all atoms in peptides are exposed to solvent to some degree; there is no sharp distinction between surface and interior groups. Secondly, peptides adopt conformations that are necessarily determined by short-range interactions. Finally, they are often conformationally mobile and their preferred conformations are strongly influenced by interactions with other molecules and by the nature of their environment. Although these characteristics of linear peptides make them difficult targets of conformation analysis, their intrinsic flexibility makes them sensitive indicators of the relative stabilities of various conformational states. The observation of one conformation for a linear peptide in solution or in crystals provides evidence for the stability of that conformation relative to other conformational states which are potentially available.

The conformations of linear peptides have been of interest to theoreticians and experimentalists for many years. The idea that secondary structures have functions related to their structural characteristics appears throughout the literature[55,56]. Recently, the discovery that peptide fragments of proteins in solution, especially in water, do, in many cases, exhibit significant conformational preferences for recognizable structures[78-85] was of great interest and significance in several fields, notably those that sought insights into the earliest events that initiate protein folding[86-90] and into the likely mechanism of induction of anti-peptide antibodies[91].

In order of the frequency in which they appear in peptides, regular secondary structures of peptides can be classified into the following classes:

Reverse Turns and Nascent Helices

Reverse turns (also known as tight turns, β bends, hairpin bends, etc.) are the simplest regular conformation formed by linear peptides. A turn is defined as a site where the polypeptide chain reverses its overall direction. In proteins, their frequent occurrence, accounting for about one-third of the total molecule[73], is responsible for the globularity of globular proteins. Turns are intrinsically polar structures with backbone groups that pack together closely and side chains that project outward. Such an array of atoms may constitute a site for molecular recognition, and indeed, the literature abounds with suggestions that turns serve as loci for receptor binding, antibody recognition, and posttranslational modification[55,56]. In addition, reverse turns probably play a key role in the initiation of protein folding and have been proposed as pivotal intermediates in the formation and denaturation of helices.[86,88,90]

Turns found in peptides can be classified into the β-turn and γ-turn, i.e., turns of four or three residues, respectively, using hard-sphere calculations and model peptides[73]. These turns may or may not be stabilized by an intra-turn hydrogen bond; in β-turns between the C=O of residue i and the NH of residue i+3, and in γ-turns between the C=O of residue i and the NH of residue i+2. The term "open" β or γ turns could be used for situations in which no hydrogen bond exists, and the ϕ, ψ angles are within 30° of the ideal. β-turns can be further classified as types I, II, III, their mirror images I', II', III', VIa(cis) and VIb(cis). γ turns can be further classified as regular and inverse γ turns[73,74]. Table 1.1 lists these turn classifications for both β and γ turns with the ϕ, ψ angles for the central residue(s)[73]. Among all these turns, type I is the most common one accounting for about 35% of the total occurrences, followed by type II and III with a 15% probability for each. The type III turn actually has repeating torsion angles and is identical with the 3_{10}-helix. For both Type III and III', the ideal ϕ, ψ

values are so close to the values for Type I and I' that they could be distinguished only in the most highly refined protein structures. Type I', II' and III' are disfavored because of steric hindrance. Type II usually requires a Gly residue in position 2. Type VI has a *cis*-proline in position 3; Type VIa has ϕ, ψ values in the α region of the Ramachandran plot for the proline, a "concave" orientation of the middle peptide and the proline ring relative to the overall curve of the turn and typically is hydrogen-bonded. Type VIb has ϕ, ψ values in the β region of the Ramachandran plot for the proline, has a "convex" orientation of the middle peptide and the proline ring and is usually not hydrogen-bonded.

The term "nascent helix" has been used to describe an ensemble of turn-like structures over several adjacent residues of a peptide[80]. Nascent helix structures typically occur in peptide fragments derived from helical segments of protein.

Helices

Helical conformations have been demonstrated by a variety of techniques in several linear peptides[81]. The formation of helices in peptides has also been implicated in their ability to induce membrane fusion and in the mechanism of action of signal peptides[81].

The α-helix has 3.6 residues per turn, with a hydrogen bond between the CO of residue i and the NH of residue i + 4. The closed loop formed by one of these hydrogen bonds and intervening stretch of backbone contains 13 atoms (including hydrogen). In the usual nomenclature for describing the basic structure of polypeptide helices, the α-helix is known as the 3.6_{13}-helix. The rise per residue along the helix axis is 1.5 Å. All α-helices are right-handed (for L-amino acids, the left-handed α-helix has a close approach between the

carbonyl oxygen and the β-carbon). With 3.6 residues per turn, side chains protrude from the α-helix at about every 100° in azimuth. The backbone conformational angles for a right-handed α-helix are approximately ϕ = -60°, ψ = -60°, which is in a favorable local conformation with a steep energy minimum, even ignoring the hydrogen bonds. α-helices are certainly the most regular pieces of structure to be found in globular proteins, but even so they show significant imperfections, such as bends in the axis of a helix.

The only other principle helical species besides the α-helix which occurs to any great extent is the 3_{10}-helix, with a three-residue repeat and a hydrogen bond to residue i+3 instead of i+4. Its backbone conformational angles are approximately ϕ = -60°, ψ = -30°, within the same energy minimum as the α-helix. However, for a long periodic structure, the 3_{10}-helix is considerably less favorable than the α-helix in both local conformational energy and in hydrogen bond configuration. Long 3_{10}-helices are very rare, but short pieces of approximate 3_{10}-helix occur fairly frequently. Two consecutive residues in 3_{10} conformation form a type III β turn, and three consecutive 3_{10}-residues forming two interlocked tight turns are also fairly commonly seen. One important location for a single turn 3_{10}-helix is at the C-terminal end of an α-helix.

The generally regular, repeating conformation in the α-helix places all of the charge dipoles of the peptides pointing in the same direction along the helix axis (positive toward the N-terminal end). It has been shown that the overall effect is indeed a significant net dipole for the helix, in spite of shielding effects, The helix dipole may contribute to the binding of charged species to the protein.

β-Sheets

Few well-characterized examples of β-sheet formation by short linear peptides are found because sheets are generally formed by association.

Principally, a β sheet is made up of almost fully extended strands, with φ, ψ angles which fall within the wide, shallow energy minima in the upper left quadrant of the Ramachandran plot (φ, -180° to 0°, ψ, 0° to 180°). β strands can interact in either parallel or antiparallel orientation, and each of the two forms has a distinctive pattern of hydrogen bonding. The antiparallel sheet has hydrogen bonds perpendicular to the strands, narrowly spaced bond pairs alternate with widely spaced pairs. Looking from the N- to C- terminal direction along the strand, when the side chain points up, the narrow pair of H - bonds will point to the right (C=O to NH). A parallel sheet has evenly spaced hydrogen bonds which angle across between the strands. Within a β sheet, as within an α-helix, all possible backbone hydrogen bonds are found. In both parallel and antiparallel β sheets, the side chains of adjacent amino acids along each strand alternate above and below the sheet, while those on the same side of the sheet are quite close together.

Spectroscopic Methods to Detect Secondary Structures in Small Peptides

Most, if not all, short linear peptides are present in solution as an ensemble of rapidly interconverting conformations. In favorable cases, the population of one or more folded conformers in the ensemble is high enough for detection by spectroscopic methods. Several spectroscopic techniques have been used to elucidate the solution conformations of linear peptides, including circular dichroism (CD), vibrational spectroscopy (Raman, Fourier transform infrared (FTIR)), and nuclear magnetic resonance (NMR). In this thesis, NMR and CD were used to study the solution conformation of several peptide fragments in methanol/water solvent.

NMR Spectroscopy

Nuclear magnetic resonance (NMR) is the most useful method for peptide conformational analysis in solution. NMR yields information about chemical environments of individual nuclei (chemical shifts), geometric relationship between nuclei (coupling constants), distances between nuclei (nuclear Overhauser enhancements (NOEs)), accessibility of amide protons (exchange kinetics and sensitivity of resonance positions and linewidths to temperature, solvent, or paramagnetic probes), and dynamics of nuclei (relaxation times). More importantly, this wealth of information is site specific, so that properties of individual residues can be studied[92]. When applied appropriately, NMR provides an exceptionally powerful technique for determining the structure of the dominant conformers adopted by a peptide in solution.

The data for conformational analysis can be collected using one-, two- or three-dimensional NMR experiments[92-96]. For small linear peptides without isotope labeling, the most commonly used experiments are homonuclear two-dimensional COSY, DQCOSY, TOCSY, NOESY and ROESY experiments, which will be discussed in detail. Figure 1.3 shows the pulse schemes for these experiments.

A 2D NMR experiment generally includes four successive time periods, preparation, evolution, mixing and detection. Preparation, evolution and detection are mandatory, and most 2D NMR experiments include a mixing period or a mixing pulse. The preparation period usually consists of a delay time, during which thermal equilibrium or steady state is attained, followed by one or more radiofrequency (rf) pulses to create the desired coherence. During the evolution period the coherence of the system is frequency-labelled in a manner which depends on the length of the evolution period. The mixing period may include one or more rf pulses and delay intervals. During the mixing period, information

is transferred between spins by some physical process, so that connections between nuclei labelled in the evolution period and nuclei detected during detection period are established. During the detection period the system evolves further and resulting free induction decay (FID) is recorded.

In the two-dimensional homonuclear correlated spectroscopy experiment (COSY)[97], mixing is achieved by a 90° pulse separating the evolution and the detection periods[98]. The COSY experiment produces correlation maps that display connectivities due to scalar spin-spin coupling and thus provides information on the proximity of nuclei connected by the chemical bonds[92]. For amino acid residues in peptides and proteins, COSY experiments can usually provide correlation between protons within each residue and the coupling constants between the NH and α-protons. COSY is normally performed using phase cycling which gives quadrature detection and phase sensitive presentation in both dimensions[99].

In DQCOSY (double quantum filtered COSY) experiments[100], the first 90° pulse creates single quantum coherence which is subsequently frequency labeled during the evolution period t_1, as in conventional COSY. DQCOSY differs from conventional COSY in the coherence transfer pathway taken after the evolution period. In COSY, single quantum coherence is transferred from one spin to another; in DQCOSY the coherence transfer is through a transient state of multiple quantum coherence. The second 90° pulse excites all orders of multiple quantum coherence and the third pulse converts multiple quantum coherence into detectable single quantum coherence. The filter order is determined solely by the phase cycle of the second pulse. DQCOSY removes much of the dispersive contribution to the diagonal of the COSY spectrum and greatly attenuates the amplitudes of the diagonal peaks, which in turn reduces the amount of t_1 noise and makes it easier to analyze cross-peaks near the

diagonal. In DQCOSY, singlet resonances in the spectrum are removed and this increases the efficiency of water suppression. In addition, multiplets in DQCOSY spectra often have unique fine structures which provide increased resolving power in crowded regions plus information on the relative signs of the active couplings, therefore allowing distinction between geminal and vicinal coupling[100,101].

TOCSY, total correlation spectroscopy, was first proposed by Braunschweiler and Ernst[102]. This experiment has particular potential for protein NMR studies since it permits correlation of all protons within a scalar coupling network. Hence, a complete subspectrum can be obtained for every amino-acid, making resonance assignments considerably easier[103]. TOCSY is based on the homonuclear Hartmann-Hahn (HOHAHA) effect[104]. Magnetization transfer via the Hartmann-Hahn effect occurs only when the difference in the absolute magnitudes of the effective rf fields experienced by two coupled spins is smaller than the scalar interaction between the two spins. The effective rf field is the vector sum of the resonance offset and the applied rf field. In general, two coupled spins have different chemical shifts. One way to make the difference between their effective rf fields smaller is to supply a strong rf irradiation through composite pulse decoupling schemes, such as MLEV17[105]. In the pulse sequence shown in Figure 1.3, trim-pulses (1-2 ms) before and after the MLEV17 mixing pulses are used to give the pure phase for the diagonal. One possible artifact comes from transverse NOE effects. This effect is always positive and NOE cross-peaks are of opposite sign to the cross-peaks from Hartmann-Hahn transfer. Since NOE cross-peaks are usually weak and easy to identify, together with the fact that Hartmann-Hahn transfer is very efficient, transverse NOE effects are not a major problem[103].

The nuclear Overhauser enhancement or nuclear Overhauser effect (NOE)[106] is the fractional change in the magnetization of one spin as a consequence of perturbing the population of other spins. NOE effects are due to dipolar interactions (through-space) between different nuclei and are proportional to the inverse sixth power of the internuclear distance. In the NOESY[107,108] (2D nuclear Overhauser enhancement) experiment, the mixing period consists of two 90° pulses separated by a mixing time τ_m, during which cross-relaxation involving longitudinal magnetizations occurs. The maximum distance sampled depends on the mixing time (a mixing time of 100 ms for a protein of molecular mass 5000 yields cross-peaks from pairs of protons closer than 3Å). In an ideal situation, distances of 5Å can be detected. The length of τ_m is varied randomly by ±10% to suppress scalar-coupling effects[109]. The cross peaks in NOESY correlation maps provide evidence of dipole-dipole coupling between nuclear spins in close spatial proximity.

The ROESY experiment measures cross-relaxation rates in the rotating frame. This experiment was originally proposed by Bothner-By et al. and was termed CAMELSPIN[110], but was subsequently renamed ROESY, i.e. rotating-frame Overhauser effect spectroscopy[111]. In this experiment, a strong, on-resonance transverse radio frequency (rf) field is applied to the spins at the end of the evolution period t_1. This causes the spin magnetization to become "spin locked" parallel to the rf magnetic field. The on-resonance transverse rf field is usually achieved by a series of a pulse of tip angle and a delay. Because of their different precession frequencies during t_1, different spins have different amounts of magnetization aligned along the rf field. Net transfer of transverse magnetization can then take place between spatially proximal spins. As with NOESY, the ROESY experiment allows all pairs of spatially proximal spins to be detected simultaneously.

In ROESY experiments, transverse magnetization undergoes cross-relaxation and the rate of cross-relaxation therefore depends on spin-spin relaxation processes. In contrast, in NOESY experiments longitudinal magnetization undergoes cross-relaxation and the rate of cross-relaxation therefore depends on spin-lattice relaxation processes. Consequently, cross-relaxation in the rotating frame has a different dependence on molecular motion than cross-relaxation in the laboratory frame. In the rotating frame, the cross-relaxation rate is always positive. In contrast, in the laboratory frame the cross-relaxation rate changes from negative to positive with increasing molecular size. For medium size molecules, the cross-relaxation rate in the laboratory frame can be close to zero, resulting in unobservable NOEs. This means that the ROESY experiment is particularly attractive for moderately sized molecules like small linear peptides. For bigger molecules, cross-relaxation in the laboratory frame is an energy-conserving process and this leads to substantial spin diffusion. In contrast, for slow molecular motion cross-relaxation in the rotating frame is tempered by dipolar relaxation, therefore attenuating spin diffusion which means more accurate measurements of cross-relaxation rates and better descriptions of molecular structure. Finally, combined with NOESY, ROESY provides an attractive method to study molecular motion. Measuring cross-relaxation rates both in the rotating frame and in the laboratory frame may allow the nature of the molecular motion to be defined[112].

It is possible that some artifacts could appear in the ROESY experiment[112]. Among them are ones mediated by chemical exchange and the most important coherence transfer mechanism via J-coupling, i.e., COSY type transfer and homonuclear Hartmann-Hahn transfer. Multistep magnetization transfer is also possible. The cross-peaks resulting from COSY type transfer can usually be recognized by their anti-phase character. The cross-peaks

resulting from chemical exchange and Hartmann-Hahn transfer show a positive absorption, while the real NOE cross-peaks show negative absorptions. To reduce the Hartmann-Hahn transfer in the ROESY experiment, a small angle pulse (30°) for the transverse rf field and a moderate spin-lock field strength should be used. Hard 90° pulses immediately before and after the spin-lock are also used to reduce the resonance offset effects.

A conformational analysis of peptides or proteins by NMR starts with sequence-specific resonance assignments[92]. The most successful and rigorous procedure for sequential assignments is that proposed by Wüthrich and co-workers[92], based on the observation of sequential NOESY or ROESY cross-peaks. The overall strategy can be divided into two parts; (1) identification of spin systems using mainly H-H spin-spin connectivities, which can be done by using COSY, DQCOSY and most efficiently by TOCSY experiments; and (2) sequential assignments of these spin systems based on the observed NOEs between adjacent residues, starting with unique residues and unique pairs of residues for which the NOEs give the highest probability of being sequential. The short distances that are important for sequence-specific assignments are distances between the α proton of one residue and the NH proton of the next residue, the distance between the β proton of one residue and the NH proton of the next residue, and the distance between sequential NH protons. Once sequence-specific resonance assignments have been obtained, NOEs can be attributed to specified locations in the sequence. Based on this specific NOE information, combined with further NMR measurements, such as chemical shifts, coupling constants, amide-proton exchange rates and temperature dependence coefficients of amide protons, the conformation of the molecule can be analyzed in detail.

Chemical shifts do not give conformational information directly. Indirectly, their values do have some implications for conformation analysis, such as the involvement in hydrogen bonds and closeness to an aromatic center, when the chemical shift values are quite different from their average values. In some cases, chemical shifts can serve to monitor conformational changes in the molecule when the environment changes. Recently, with the increase in chemical shift data for protein molecules, correlation between chemical shifts and secondary structures of α-helix and β-sheets in proteins has been demonstrated[113,114].

^1H-^1H homonuclear coupling constants can be measured from 1D or 2D NMR experiments. Among 2D experiments, COSY and DQCOSY are the most common ones. When the coupling constants are comparable to the linewidth, care must be taken since the antiphase cross peak separation does not reflect the coupling constant[93]. In this case, other experiments, such as ECOSY[115] and P.E.COSY[116], or direct simulation of cross peaks[117] may be used. Homonuclear coupling constants can provide valuable information for the conformation by means of the well-known Karplus equation[118] using the Karplus constants obtained from the globular protein BPTI (Basic pancreatic trypsin inhibitor)[119], combined with NOE information. In principle, the backbone conformations of peptides are defined by the torsion angles ϕ, ψ and ω of each residue, which, in turn, can be derived from the Hα-NH coupling constants and the strength of the nuclear Overhauser enhancements. The torsion angles ϕ, describing the rotation about the N-Cα bond, can be estimated by using the Karplus equation directly. In general, as many as four different dihedral angles may correlate with a given coupling constant. In peptides and proteins, however, ϕ angles are found to be in the range from -30° to -180°[73] and, for coupling constants smaller than 6.0 Hz, they fall into the -50° to -90° range. In addition,

NH(i)-Hβ(i) NOE information can often provide the necessary information to distinguish between φ-values around -80° and around -160°, both related to a 7.0 Hz coupling constant. Assuming that only the three staggered side-chain rotamers are energetically favorable, and allowing a deviation of +/- 30° from the ideal χ1 (χ1 is the dihedral angle between C^{α}(i) and C^{β}(i) atoms) values of -60°, +60° and 180°, the distance range between NH(i) and any of the two β-protons (i) is found to be between 2.2 and 2.8 Å for φ = -80° and between 3.0 and 3.6 Å for φ =-160°[120]. In the first case, medium to strong nuclear Overhauser effects are to expected, whereas in the second case only weak interactions should occur. The ψ angles cannot be derived from homonuclear coupling constants[121], but they often can be estimated satisfactorily from the Hα(i)-NH(i+1) NOEs. Assuming trans peptide bonds (ω = 180°), the distance between Hα(i) and NH(i+1) is minimal, and hence the NOE maximal, for ψ = +120°. Within deviations of +/- 30°, the change in distance is very small[120] and clearly below 2.5 Å, which is generally considered as the upper limit for NOEs to be labeled as "strong"[92]. Medium and, in particular, weak NOEs allow less definitive conclusions about the ψ angles, and should therefore be considered as approximations. The peptide bond angle ω is, except for peptides containing proline, considered to be 180° (trans). The conformation of side-chains is analyzed in terms of side-chain rotamer populations[122] for which clearly resolved Hα-Hβ and Hα-Hβ' coupling constants are a prerequisite.

The most important experimental structural information obtained from the NMR studies is the NOE between two spins from NOESY or ROESY experiments since distance information can be gained from NOEs. For a pair of spins, the buildup of crosspeak intensity for a short mixing time is approximately linear with a slope proportional to τ_c/r^6 (τ_c is the overall molecular tumbling correlation time and r the distance between two nuclei). In principle, a more

quantitative estimate of the distance can be made by using the buildup curve[92], crosspeak intensity vs. mixing time. Since the correlation time is not known in general, distances between spins may be calculated from ratios of crosspeak intensities, using one crosspeak for which the distance is known (i.e. geminal pair), provided the effective correlation time is the same for all pairs of spins. If spin diffusion occurs, or different parts of the molecule have different correlation times, more complicated procedures, such as complete relaxation matrix analysis, must be used to get correct distances[123].

Temperature dependence coefficients of amide protons, $\Delta\delta/K$, can be determined from studies of the chemical shifts of amide protons at different temperatures. Smaller values of $\Delta\delta/K$ usually indicate that the amide protons are involved in hydrogen bonds or are otherwise inaccessible to solvent molecules. For large molecules, exchange rates of amide protons can be measured from a series of spectra taken in a completely deuterated solvent. Slow exchange rates are an indication of hydrogen bond involvement and of less structure fluctuation. Relaxation measurements (both T_1 and T_2) can be used to detect the presence of other motion, such as chemical and conformational exchange[92].

In principle, most of the NMR information can be transferred into geometric constraints over the conformation for the molecule studied. Based on these constraints, the molecular conformations can be generated by using distance geometry and molecular mechanics/dynamics calculations[124,125]. However, the general rules above can only be applied when the studied peptide possesses a relatively well-defined conformation or rigid conformation, which might be the case in some protein molecules. In the case of small linear peptides, the peptides possess an unlimited number of conformations and conformational interconversions not requiring peptide bond rotations occur in a time scale of nanoseconds. Since NMR is a method with a time scale of

seconds to hundredths of seconds for the usual NMR measurements, all measurable parameters for small linear peptides are therefore population-weighted averages of the conformational states available to the peptide, which precludes numerically precise interpretation of the data. Nevertheless, secondary structures can be identified by some characteristic distances, coupling constants or both for that specific secondary structure. Some important characteristic distances and coupling constants for various secondary structures are listed in Table 1.2[92]. Wright[126] summarized a list of accepted criteria, some of which are considered more definitive than others, that can be used to define the secondary structures of small linear peptides in solution.

The criteria for turn formation include: (a) $d_{NN}(i, i+1)$ NOE connectivities between residues 3 and 4 of the turn (type I turns also have a $d_{NN}(2,3)$ connectivity); (b) a $d_{\alpha N}(2,4)$ NOE connectivity; (c) appropriate $^3J_{HN\alpha}$ coupling constants (not always seen); (d) a low temperature coefficient for the NH of residue 4 of the turn, indicating hydrogen bonding (not always seen). Such a hydrogen bond is not a requirement for turn formation, at least not in proteins[127]; (e) characteristic CD or vibrational spectra.

The criteria for helix formation include: (a) a series of relatively strong $d_{NN}(i, i+1)$ NOE connectivities, usually accompanied by a diminution in the magnitude of the $d_{\alpha N}(i, i+1)$ NOE; (b) reduced $^3J_{HN\alpha}$ coupling constants; (c) medium-range $d_{\alpha N}(i, i+3)$, $d_{NN}(i, i+2)$, $d_{\alpha\beta}(i, i+3)$, and $d_{\alpha\beta}(i, i+4)$ NOEs; (d) a pronounced negative ellipticity at 223 nm in the CD spectrum (the wavelength of the minimum of the 200 to 208 nm band is also a good indicator of helical conformation; (e) characteristic vibrational frequency bands in Raman or FTIR spectra.

The criteria for β-sheets are characteristic interstrand $d_{\alpha N}(i, j)$, $d_{NN}(i, j)$ and $d_{\alpha\alpha}(i, j)$ NOEs.

Circular Dichroism

It is well known that molecules such as peptides and proteins containing asymmetric carbon atoms are optically active and exhibit the phenomenon of optical rotatory dispersion (ORD), which is the measurement of the wavelength dependence of the "dispersion" of optical rotation, and of the circular dichroism (CD) caused by different interactions of left-handed circularly polarized light (LHCPL) and right-handed circularly polarized light (RHCPL) with chiral molecules[128]. A solution of optically active molecules absorbs or refracts LHCPL and RHCPL to different extents. If ε_L and ε_R are the molar absorption coefficients for LHCPL and RHCPL, the molar circular dichroism is defined as $\Delta\varepsilon = \varepsilon_L - \varepsilon_R$, which naturally gives rise to the observable property of ellipticity. If n_L and n_R are refractive indices for LHCPL and RHCPL, the circular birefringence or optical rotation is a measure of $(n_L - n_R)$, which gives rise to optical rotation. Because of the close relationship between absorbance and refraction, CD and ORD properties are interrelated by a Kronig-Kramers transform. However, because of the greater sensitivity of measurements of the circular dichroic absorbance than of the ellipticity, CD has replaced ORD for the conformational analysis of proteins and peptides[128,129].

Circular Dichroism (CD) is a very sensitive technique. In practice, differences of as little as one part in 10^6 in absorbances are sufficient to generate readily observable effects in CD. Probably the most extensively used CD technique for the exploration of peptide and protein conformation in solution is the CD spectrum in the far-ultraviolet (UV) region (170 nm to 240 nm), where the backbone amide chromophore absorbs. Since the arrangement of peptide bond chromophores with respect to each other and to asymmetric fields in the molecule will vary with different conformations, far UV circular dichroism is

exquisitely sensitive to conformation. CD also has a very fast time scale (10^{-15} sec.) and is able to detect an individual conformation for the peptides which undergoes fast exchange with the others in term of the time scale in NMR. The disadvantage for CD is that spectra of peptides and proteins are composed of broad overlapping bands arising from each amide chromophore of the molecule. Consequently, detailed interpretation of CD spectra for molecules larger than di- or tripeptides is difficult and information provided is not site-specific. Furthermore, several different conformations could yield similar CD spectra.

Fortunately, the characteristic CD spectra of regular elements of secondary structure present in proteins and large polypeptides are well understood from experimental and theoretical work[128-131]. An ideal α-helix usually shows a typical double minimum shape with minima at 222 and 208-210 nm of approximately equal intensity, a maximum at 191-193 nm, which is about twice as intense as those negative bands, and a positive shoulder near 175 nm[132]. For a less than ideal α-helix, the intensities of the three major bands could be quite different from the ones of an ideal α-helix[131]. The 3_{10} helix is predicted to have a spectrum similar to that of the α-helix. The length of the helix is also an important factor for the overall shape of the CD spectrum. It is predicted that a minimum of 2 turns (~ 10 residues) are necessary to provide an α-helix-like CD spectrum, while 50 residues provide an accurate model of an infinite helix[131]. β-sheets are more diverse than helices; their CD spectra depend on the polarity (parallel versus antiparallel), the number of strands, and the number of residues per strand and their twists. A β-sheet could show a negative band above 200 nm, a positive one near 200 nm and another negative band below 170 nm[133,134]. Parallel and antiparallel β-sheets have been predicted to have qualitatively similar CD spectra[133]. Among the many variants of the β-sheet, the positions of the bands can be shifted by several nanometers

and their magnitudes varied. The CD spectrum of the unordered form generally shows a strong negative band near 200 nm and a very weak band around 220 nm, which can be either a positive band or a negative shoulder. The magnitude of the 200 nm band can vary with the molecules and the conditions[135]. The situation for the CD spectra of β-turns is even worse, because β-turns have more variants. Theoretical analysis predicts that type I, II and III β-turns, which account for about 80% of β-turns in proteins[136], should have similar CD spectra[137]. The most common CD spectrum for turns shows a weak negative band between 220-230 nm, a positive band between 200 and 210 nm and a strong negative band between 180 and 190 nm. However, CD spectra of some β-turns may be quite different.

More importantly, the chiroptical properties of various conformations in a molecule are assumed to be additive. The contributions due to non-peptide chromophores below 250 nm are frequently neglected, This may be justifiable if aromatic groups and disulfide bonds account for less than 10% of the residues. The CD spectrum of a protein can be deconvoluted in terms of contributions from different components, such as α-helix, β-sheet, reverse turn, and so-called random coil, provided that the reference spectrum for each component can be evaluated. Reference spectra can be obtained from model compounds[129,133], or calculated from CD spectra of proteins of known secondary structure[130,138,139]. With such approaches, CD data can give statistically significant estimates of many classes of secondary structure, if data are measured from 250 nm to 184 nm. If data extending only to 200 nm are used, the presence of α-helix is generally the only one that can be determined with confidence[130,140]. In one case, correlation coefficients between X-ray structure and predicted structure for CD data measured to 178 nm are 0.97 for α-helix, 0.76 for β-sheet, 0.49 for β-turns and 0.86 for other structures[141].

In principle, the same methods used to identify and deconvolute protein CD spectra can be used to study small linear peptides. A population of secondary structure will be mostly readily detected by CD in a small peptide because the ellipticity is measured per residue and would be diluted out in a longer peptides by the signals from the rest of peptide. However, care must be taken because of the dynamic properties of small peptides. In peptides, secondary structures are less ideal and more flexible, so that their CD spectra could be quite different from those of ideal secondary structures[81,82]. Deconvolution could be very difficult, because of the absence of reference spectra. Especially, the reference spectra for β-turns are difficult to obtain and have relatively low ellipticities compared to α-helix or β- structure, which means that decomposing a complex spectrum into various structural features including turns involves large errors.

References

1. Bornstein, P. & Sage, H. (1980), *Ann. Rev. Biochem.* **49**, 954-1003.
2. Miller, E. J. & Gray, S. (1982), *Methods in Enzymology* **82**, 3-32.
3. Ramachandran, G. N. & Ramakrishnan, C. (1976), *Biochemistry of Collagen*, Ramachandran, G. N. & Reddi, A. H. Eds, New York and London, Plenum Press.
4. Eyre, D. R., Paz, M. A. & Gallop, P. M. (1984), *Annu. Rev. Biochem.* **53**, 717-748.
5. Weiss, J. B. (1976), *Int. Rev. Connect. Tissue Res.* **7**, 101-157.
6. Pope, F. M. & Nicholls, A. C. (1984), *Molecular Medicine*, Vol. **I**, Malcolm, A. D. B., Ed., IRL Press, Oxford, 117-175.
7. Light, N. D. & Bailey, A. J. (1979), *Fibrous Proteins: Scientific, Industrial and Medical Aspects*, Vol. I, Parry, D. A. D. & Cramer, L. K. Eds., Academic Press, London, 151-177.
8. Parry, D. A. D., Graig, A. S. & Barnes, G. R. G. (1978), *Proc. Royal Soc. (London)* **203**, 305-321.
9. Parry, D. A. D. (1984), *Ultrastructure of the Connective Tissue Matrix*, Ruggeri, A. & Motta, P. M. Eds., Martinus-Nijhoff, Boston, 34-64.
10. Fietzek, P. P., Kühn, K. (1976), *Int. Rev. Connect. Tissue Res.* **7**, 1-60.
11. Glanville, R. W. & Kühn, K. (1979), *Fibrous Proteins: Scientific, Industrial and Medical Aspects*, Vol. I, Parry, D. A. D. & Cramer, L. K. Eds., Academic Press, London, 133-150.
12. Chapman, J. A. & Hulmes, D. J. S. (1984), *Ultrastructure of the Connective Tissue Matrix*, Ruggeri, A. & Motta, P. M. Eds., Martinus-Nijhoff, Boston, 1-33.
13. Su, M. W., Lee, B., Ramirez, F., Machado, M. & Horton, W. (1989), *Nucleic Acids Res.* **17**, 9473.

14. Miller, E. J. (1984), *Extracellular Matrix Biochemistry*, Elsevier, New York, 41-81.

15. Lapiere, Ch. M., Nusgens, B. & Pierard, G. E. (1977), *Connective Tissue Research* **5**, 21-29.

16. Keene, D. R., Sakai, L. Y., Bächinger, H. P. & Burgeson, R. E. (1987), *J. Cell. Biol.* **105**, 2393-2402.

17. Fietzek, P. P., Allmann, H., Rauterberg, J., Henkel, W., Wachter, E. & Kuhn, K. (1979), *Hoppe-Seyler's Z. Physiol. Chem.* **360**, 809-820.

18. Chu, M.-L., Weil, D. DeWet, W., Bernard, M, Sippola, M. & Ramirez, F. (1985), *J. Biol. Chem.* **260**, 4357-4363.

19. Yamada, Y., Kühn, K. & de Crombrugghe, B. (1983), *Nucleic Acids Res.* **11**, 2733-2744.

20. Gross, J. & Kirt, D. (1958), *J. Biol. Chem.* **233**, 355-360.

21. Williams, B. R., Gelman, R. A.. Poppke, D. C. & Piez, K. A. (1978), *J. Bio. Chem.* **255**, 6578-6585.

22. Wood, G. C. & Keech, M. K. (1960), *Biochem. J.* **75**, 588-598.

23. Wood, G. C. (1964), *Int. Rev. Connect. Tissue Res.* **2**, 1-31.

24. Gross, J., Highberger, J. H. & Schmitt, F. O. (1955), *Proc. Natl. Acad. Sci. U.S.A.* **41**, 1-7.

25. Jackson, D. S. & Fessler, J. H. (1955), *Nature* (London) **176**, 69-70.

26. Yuan, L. & Veis, A. (1973), *Biopolymer* **12**, 1437-1444.

27. Bernengo, J. C., Herbage, D., Marion, C. & Roux, B. (1978), *Biochim. Biophys. Acta* **532**, 305-314.

28. Comper, W. D. & Veis, A. (1977), *Biopolymers* **16**, 2133-2142.

29. Gelman, R. A., Williams, B. R. & Piez, K. A. (1979), *J. Biol. Chem.* **254**, 180-186.

30. Leibovich, S. J. & Weiss, J. B. (1970), *Biochim. Biophys. Acta.* **214**, 445-454.

31. Helseth, D. L. Jr. & Veis, A. (1981), *J. Biol. Chem.* **256**, 7118-7128.

32. Gelman, R. A., Poppke, D. C. & Piez, K. A. (1979), *J. Biol. Chem.* **254**, 11741-11745.

33. Helseth, D. L. Jr. & Veis, A. (1979), *Fed. Proc.* **38**, 819.

34. Capaldi, M. J. & Chapman,, J. A. (1982), *Biopolymers* **21**, 2291-2313.

35. Helseth, D. L. Jr., Lechner, J. H. & Veis, A. (1979), *Biopolymers* **18**, 3005-3014.

36. Helseth, D. L. Jr. & Veis, A. (1981), *Chemistry and Biology of Mineralized Connective Tissues*, Veis, A. Ed., Elsevier/North-Holland, New York, 85-92.

37. Scott, P. G. (1986), *Biochemistry* **25**, 974-980.

38. Dion, A. S. & Myers, J. C. (1987), *J. Mol. Biol.* **193**, 127-143.

39. Jones, E. Y. & Miller, A. (1987), *Biopolymers* **26**, 463-480.

40. Otter, A., Scott, P. G. & Kotovych, G. (1988), *Biochemistry* **27**, 3560-3567.

41. Otter, A., Kotovych, G. & Scott, P. G. (1989), *Biochemistry* **28**, 8003-8010.

42. Liu, X., Scott, P. G., Otter, A. & Kotovych, G. (1990), *J. Biomol. Struct. Dyn.* **8**, 063-080.

43. Otter, A., Scott, P. G. & Kotovych, G. (1993), *Biopolymers*, in press.

44. Liu, X., Otter, A., Scott, P. G., Cann, R. & Kotovych, G., in preparation.

45. Otter, A., Scott, P. G. & Kotovych, G. (1987), *J. Am. Chem. Soc.* **109**, 6995-7001.

46. Liu, X., Scott, P. G., Otter, A. & Kotovych, G. (1992), *Biopolymers* **32**, 119-130.

47. Goodfried, T. L. & Ball, D. L. J. (1969), *J. Lab. Clin. Med.* **73**, 501-511.

48. Regoli, D., Barabe, J. C. (1980), *Pharmacol. Rev.* **32**, 1-46.

49. Farmer, S. G. & Burch R. M. (1991), *Bradykinin Antagonists, Basic and Clinical Research*, Burch, R. M. Ed., Marcel Dekker, New York, 1-31.

50. Mason, D. T. & Melmon, K. L. (1966), *J. Clin. Invest.* **45**, 1685.

51. Proud, D., Reynolds, C. J., Lacapra, S., Schotka, A., Lichtenstein, L. M. & Naclerio, R. M. (1988), *Am. Rev. Respir. Dis.* **137**, 613-616.

52. Naclerio, R. M., Proud, D., Lichtenstein, L. M., Sobotka-Kagey, A., Hemdley, J. O., Sorrentino, J. & Gwaltney, J. M. (1988), *J. Infect. Dis.* **157**, 133-142.

53. Stewart, J. M. & Vavrek, R. J. (1991), *Bradykinin Antagonists, Basic and Clinical Research*, Burch, Ronald M. Ed., Marcel Dekker, New York, 51-95.

54. Burch, R. M., Farmer, S. G. & Steranka, L. R. (1990), *Med. Res. Rev.* **33**, 237-269.

55. Rose, G. D., Gierasch, L. M. & Smith, J. A. (1985), *Adv. Protein Chem.* **37**, 1-109.

56. Degrado, W. F. (1988), *Adv. Protein Chem.* **39**, 51-124.

57. Vavrek, R. J. & Stewart, J. M. (1985), *Peptides* **6**, 161-164.

58. Schröder, E. (1970), *Bradykinin, Kallidin and Kallikrein (Handbuch Experimentellen Pharmakologie*, Erdos, E. G. Ed., Springer-Verlag, Heidelberg, Vol. **25**, 324-350.

59. Stewart, J. M. (1979), *Bradykinin, Kallidin and Kallikrein (Handbuch Experimentellen Pharmakologie*, Erdos, E. G. Ed., Springer-Verlag, Heidelberg, Vol. **25**(suppl), 227-272.

60. Stewart, J. M. & Vavrek, R. J. (1986), *Protides of the Biological Fluids*, Peeters, H. Ed., Pergamon, New York, 473-476.

61. Stewart, J. M. & Vavrek, R. J. (1986), *Kinins IV (Adv. Exp. Med. Biol. 198A*, Greenbaum, L. M. & Margolius, H. S. Ed., Plenum, New York, 537-542.

62. Vavrek, R. J. & Stewart, J. M. (1987), *Peptides 1986*, Theodoropoulos, D. Ed., deGruyter, Berlin, 655-658.

63. Vavrek, R. J. & Stewart, J. M. (1989), *Adv. Exptl. Med. Biol.* **247B**, Abe, K., Moriya, H. & Fujil, S. Ed., Plenum Press, 395-400.

64. Martorana, P. A., Kettenbach, B., Breipohl, G. Linz, W. & Scholkens, B. A. (1990), *Eur. J. Pharmacol.* **182**, 395.

65. Vavrek, R. J., Gera, L. & Stewart, J. M. (1991), *Abstract for poster presented at International Kinin Symposium*, Sept. 1991, Munich, Germany.

66. Lembeck, F., Griesbacher, T., Eckhardt, M., Henke, S., Breipohl, G. & Knolle, J. (1991), *Brit. J. Pharmacol.*, **102**, 297-304.

67. Cheronis, J. C., Whalley, E. T., Nguyen, K. T., Eubanks, S. R., Allen, L. G., Duggan, M. J., Loy, S. D., Bonham, K. A. & Blodgett, J. K. (1992), *J. Med. Chem.* **35**, 1563-1572.

68. Kyle, D. J., Martin, J. A., Farmer, S. G. & Burch, R. M. (1991), *J. Med. Chem.* **34**, 1230-1233.

69. Kyle, D. J., Hicks, R. P., Blake, P. R., Klimkowski, V. J. (1991), *Bradykinin Antagonists, Basic and Clinical Research*, Burch R. M., Ed., Marcel Dekker, New York, 131-146.

70. Lee, S. C., Russell, A. F. & Laidig, W. D. (1990), *Int. J. Peptide Protein Res.* **35,** 367-377.

71. Kyle, D. J., Martin, J. A., Burch, R. M., Carter, J. P., Lu, S., Meeker, S., Prosser, J. C., Sullivan, J. P., Togo, J., Noronha-Blob, L., Sinsko, J. A.,

Walters R. F., Whaley, L. W. & Hiner, R. N. (1991), *J. Med. Chem.* **34**, 649-2653.

72. Welsh, J. H., Zerbe, O., von Philipsborn, W. & Robinson, J. A. (1992), *FEBS Letters* **297**, 216-220.

73. Richardson, Jane S. (1981), *Adv. Protein Chem.* **34**, 167-339.

74. Schulz, G. E. & Schirmer, R. H. (1979), *Principles of Protein Structure*, Springer-Verlag, New York.

75. Creighton, Thomas E. (1984), *Proteins, Structures and Molecular Properties*, W. H. Freeman and Company.

76. Burley, S. K. & Petsko, G. A. (1988), *Adv. Protein Chem.* **39**, 125-189.

77. Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. (1963), *J. Mol. Biol.* **7**, 95-99.

78. Dyson, H. J., Cross, K. J., Houghten, R. A., Wilson, I. A., Wright, P. E. & Lerner, R. A. (1985), *Nature* **318**, 480-483.

79. Dyson, H. J., Rance, M., Houghten, R. A., Lerner, R. A. & Wright, P. E. (1988), *J. Mol. Biol.* **201**, 161-200.

80. Dyson, H. J., Rance, M., Houghten, Wright, P. E. & R. A., Lerner, R. A.(1988), *J. Mol. Biol.* **201**, 201-218.

81. Dyson, H. J., Merutka, G., Waltho, J. P., Lerner, R. A. & Wright, P. E. (1992), *J. Mol. Biol.* **226**, 795-817.

82. Dyson, H. J., Sayre, J. R., Merutka, G., Shin, H. C., Lerner, R. A. & Wright, P. E. (1992), *J. Mol. Biol.* **226**, 819-835.

83. Lark, L. R., Berzofsky, J. A. & Gierasch, L. M. (1989), *Peptide Res.* **2**, 314-321.

84. Waltho, J. P., Feher, V. A., Lerner, R. A. & Wright, P. E. (1989), *FEBS Letters.* **250**, 400-404.

85. Chandrasekhar, K., Profy, A. T. & Dyson, H. J. (1991), *Biochemistry* **30**, 9187-9194.

86. Kim, P. S. & Baldwin, R. L. (1990), *Annu. Rev. Biochem.* **59**, 631-660.

87. Fischer, G. & Schmid, F. X. (1990), *Biochemistry* **29**, 2205-2212.

88. Jaenicke, R. (1991), *Biochemistry* **30**, 3147-3161.

89. Ptitsyn, O. B. (1991), *FEBS Letters* **285**, 176-181.

90. Wright, P. E., Dyson, H. J. & Lerner, R. A. (1988), *Biochemistry* **27**, 7167-7175.

91. Dyson, H. J., Lerner, R. A. & Wright, P. E. (1988), *Annu. Rev. Biophys. Biophys Chem.* **17**, 305-324.

92. Wüthrich, K. (1986), *NMR of Proteins and Nucleic Acids*, New York, Wiley.

93. Derome, A. E. (1987), *Modern NMR Techniques for Chemistry*, Pergamon Press, New York.

94. Turner, D. L. (1985), *Prog. Nucl. Magn. Reson. Spectrosc.* **17**, 281-358.

95. Markley, J. L. (1989), *Methods in Enzymology* **176**, 12-64.

96. Clore, G. M. & Gronenborn, A. M. (1991), *Prog. Nucl. Magn. Reson. Spectrosc.* **23**, 43-92.

97. Nagayama, K., Kumar, A., Wüthrich, K. & Ernst, R. R. (1980), *J. Magn. Reson.* **40**, 321-334.

98. Bax, A. & Freeman, R. (1981), *J. Magn. Reson.* **44**, 542-561

99. Wider, G., Macura, S., Kumar, A., Ernst, R. R. & Wüthrich, K. (1984), *J. Magn. Reson.* **56**, 207-234.

100. Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R. & Wüthrich, K. (1983), *Biochem. Biophys. Res. Commun.* **117**, 479-485.

101. Rance, M., Chazin, W. J., Dalvit, C. & Wright, P. E. (1989), *Methods in Enzymology* **176**, 114-134.

102. Braunschweiler, L. & Ernst, R. R. (1983), *J. Magn. Reson.* **53**, 521-528.

103. Bax, A. (1989), *Methods in Enzymology* **176**, 151-168.

104. Müller, L. & Ernst, R. R. (1979), *Mol. Phys.* **38**, 963-995.

105. Bax, A. & Davis, D. G. (1985), *J. Magn. Reson.* **65**, 355-360.

106. Noggle, J. H. & Schirmer, R. E. (1971), *The Nuclear Overhauser Effects*, Academic, New York.

107. Griesinger, Ch. & Ernst, R. R. (1987), *J. Magn. Reson.* **75**, 261-271.

108. Jeener, J., Meier, R., Bachmann, P. & Ernst, R. R. (1979), *J. Chem. Phys.* **71**, 4546-4553.

109 Macura, S., Wüthrich, K. & Ernst, R. R. (1982), *J. Magn. Reson.* **46**, 269-283.

110. Bothner-By, A. A., Stephens, R. L., Lee, J., Warren, C. D. & Jeanloz, R. W. (1984), *J. Am. Chem. Soc.* **106**, 811-813.

111. Bax, A. & Davis, D. G. (1985), *J. Magn. Reson.* **63**, 207-213.

112. Brown, L. R. & Farmer II, B. T. (1989), *Methods in Enzymology* **176**, 199-216.

113. Wishart, D. S., Sykes, B. D. & Richards, F. M. (1991), *J. Mol. Biol.* **222**, 311-333.

114. Wishart, D. S., Sykes, B. D. & Richards, F. M. (1992), *Biochemistry* **31**, 1647-1651.

115. Griesinger, C., Sørensen, O. W. & Ernst, R. R. (1985), *J. Am. Chem. Soc.* **107**, 6394-6396.

116. Mueller, L. (1987), *J. Magn. Reson.* **72**, 191-196.

117. Widmer, H. & Wüthrich, K. (1987), *J. Magn. Reson.* **74**, 316-336.

118. Karplus, M. J. (1963), *J. Am. Chem. Soc.* **85**, 2870-2871.

119. Pardi, A.; Billeter, M.; Wüthrich, K. (1984), *J. Mol. Biol.* **180**, 741-751.

120. Leach, S. J.; Némethy, G.; Scheraga, H. A. (1977), *Biochem. Biophys. Res. Commun.* **75**, 207-215.

121. Bystrov, V. F. (1976), *Prog. Nucl. Mag. Reson. Spectrosc.* **10**, 41-81.

122. Jardetzky, O.; Roberts, G. C. K. (1981), *NMR in Molecular Biology*, Academic Press, New York.

123. Borgias, B. A. & James, T. L. (1989), *Methods in Enzymology* **176**, 169-183.

124. Kuntz, I. D., Thomason, J. F. & Oshiro, C. M. (1989), *Methods in Enzymology* **177**, 159-204.

125. Scheek, R. M., van Gunsteren, W. M. & Kaptein, R. (1989), *Methods in Enzymology* **177**, 204-218.

126. Dyson, H. J. & Wright, P. E. (1991), *Annu. Rev. Biophys. Biophys. Chem.* **20**, 519-538.

127. Chou, P. Y. and Fasman, G. D. (1977), *J. Mol. Biol.* 115, 135-175.

128. Bayley, P. (1980), *An Introduction to Spectroscopy for Biochemists*, Brown, S. B. Ed., Academic Press. New York, 148-234.

129. Yang, J. T., Wu, C-S. C. & Martinez, H. M. (1988), *Methods in Enzymology* **130**, 208-269.

130. Johnson, Jr., W. C. (1990), *Proteins* **7**, 205-214.

131. Manning, M. C. & Woody, R. W. (1991), *Biopolymers* **31**, 569-586.

132. Johnson, W. C., Jr. & Tinoco, I., Jr. (1972), *J. Am. Chem. Soc.* **94**, 4389-4390.

133. Chang, C. T., Wu, C-S, C. & Yang, J. T. (1978), *Anal. Biochem.* **91**, 13-31.

134. Woody, R. W. (1969), *Biopolymers* **8**, 669-683.

135. Woody, R. W. (1977), *J. Polymer Sci. Macromol. Rev.* **12**, 181-320.

136. Chou, P. Y. & Fasman, G. D. (1979), *Biophys. J.* **26**, 367-383.

137. Woody, R. W. (1974), *Peptides, Polypeptides and Proteins*, Blout, E. R., Bovey, F. A., Goodman, M. & Lotan, N. Ed., Wiley, New York, 338-350.

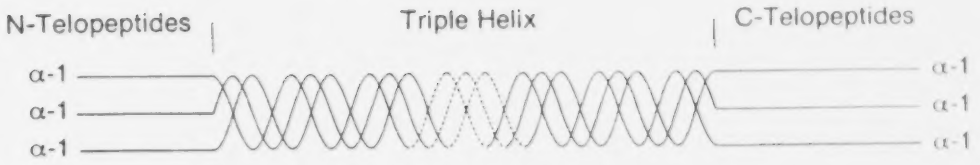
138. Perczel, A., Park, K. & Fasman, G. D. (1992), *Anal. Biochem.* **203**, 83-93.

139. Perczel, A., Hollósi, M., Tusnády, G. & Fasman, G. D. (1991), *Protein Engineering* **4**, 669-679.

140. Venyaminov, S. Y., Baikalov, I. A., Wu, C-S, C. & Yang, J. T. (1978), *Anal. Biochem.* **198**, 250-255.

141. Manavalan, P. & Johnson, W. C., Jr. (1987), *Anal. Biochem.* **167**, 76-85.

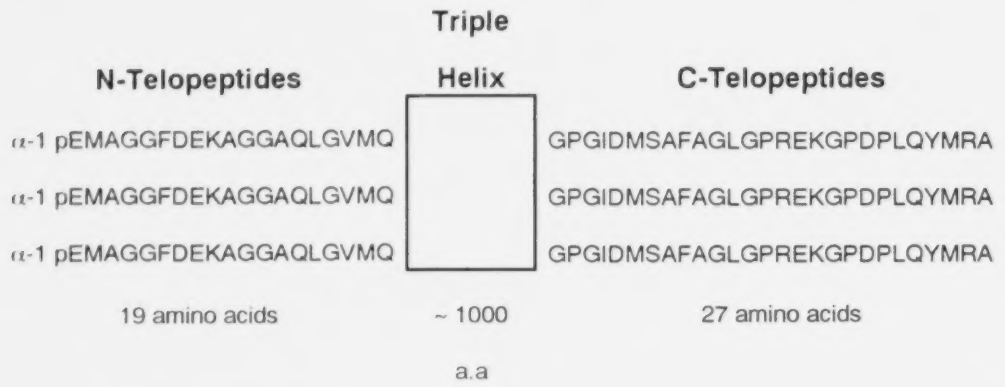
(a). Schematic Representation of Type II Collagen Structure



(b). Type I Collagen in Bovine and its Telopeptides

	N-Telopeptides	Triple Helix	C-Telopeptides
α-1	pELSYGYDEKSTGISVP		SGGYDLSFLPQPPQEKAHDGGRYY
α-1	pELSYGYDEKSTGISVP		SGGYDLSFLPQPPQEKAHDGGRYY
α-2	pEFDAKGGGP		SGGYEF
	16 amino acids	~ 1000 a.a	24 amino acids

(c) Type II Collagen in Human and its Telopeptides



(d) Type III Collagen in Bovine and its Telopeptides

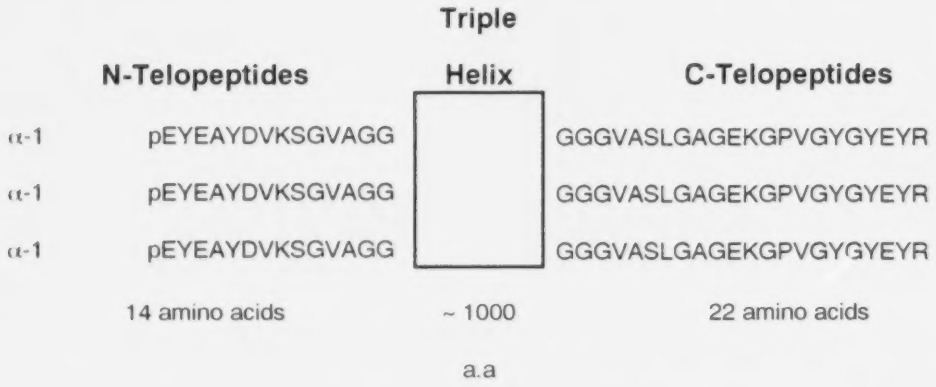


Figure 1.1 Schematic representations of a Type II collagen monomer structure and telopeptide sequences in Type I, Type II and Type III collagens.

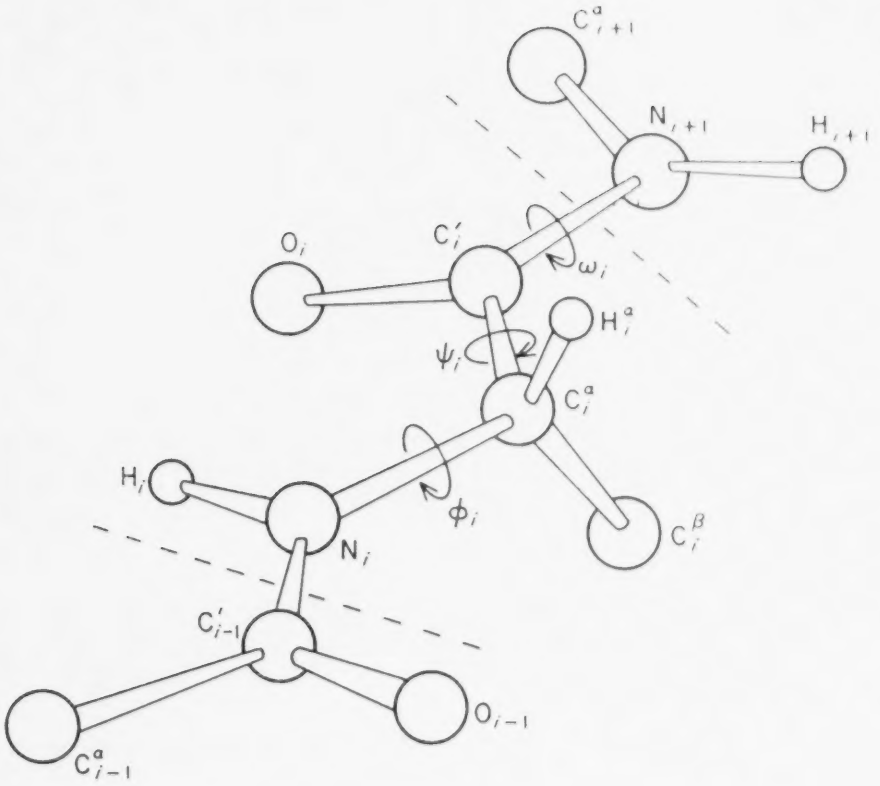


Figure 1.2 Standard nomenclature for the atoms and the torsion angles along a polypeptide chain.

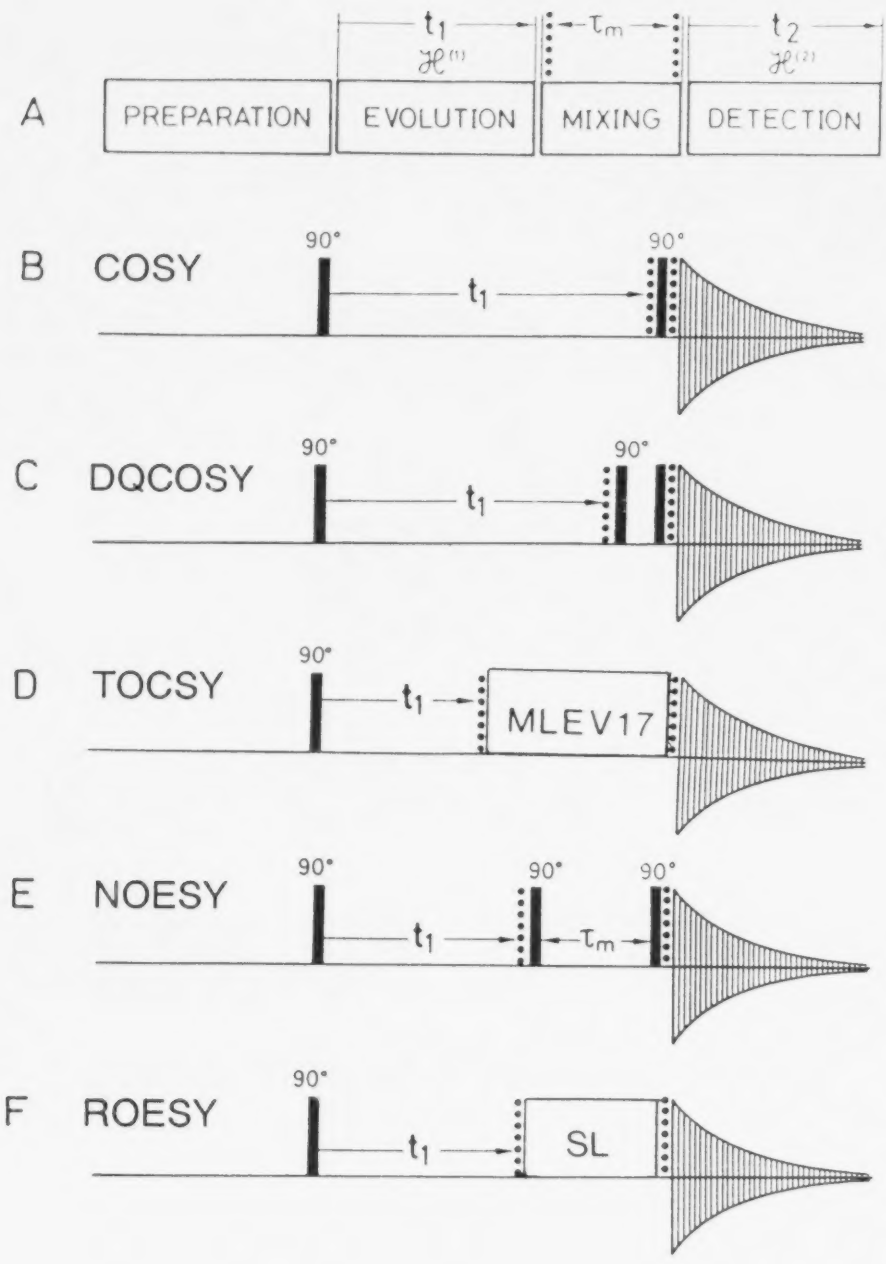


Figure 1.3 Pulse sequences for some 2D NMR experiments.

Table 1.1. Dihedral Angles of Hydrogen Bonded β and γ Turns:

	i + 1		i + 2	
β-Turn	φ	ψ	φ	ψ
Type I	-60	-30	-90	0
Type I'	60	30	90	0
Type II	-60	120	80	0
Type II'	60	-120	-80	0
Type III	-60	-30	-60	-30
Type III'	60	30	60	30
Type VIa (cis)	-60	120	-90	0
Type VIb(cis)	-120	120	-60	0
γ-turn	70 to 85	-60 to -70		
Inverse γ–turn	-70 to -85	60 to 70		

Table 1.2. Characteristic Short Distances (Å) and Coupling Constants (Hz) in Some Regular Secondary Structures

	α-helix	3_{10}	β-sheet	Type I[a]	Type II[a]
$d_{\alpha N}$	3.5	3.4	2.2	3.4	2.2
				3.2	3.2
$d_{\alpha N}(i,i+2)$		3.8		3.6	3.3
$d_{\alpha N}(i,i+3)$	3.4	3.3			
d_{NN}	2.8	2.6		2.6	
				2.4	2.4
$d_{NN}(i,i+2)$				3.8	
$d_{\beta N}$[b]	2.5-4.1	2.9-4.4	3.2-4.7	2.9-4.4	3.6-4.6
				3.6-4.6	3.6-4.6
$d_{\alpha\beta}(i,i+3)$[b]	2.5-4.4	3.1-5.1			
Coupling	4	4		4[c]	4,5[d]
Constants					

a) For the turns, the first two numbers apply to the distance between residues 2 and 3, the second to the distances between residues 3 and 4.

b) The ranges given correspond to the distances adopted by a β-methine proton if χ_1 is varied between -180 and 180°.

c) For residue 2 in the turn.

d) For residues 2 and 3, respectively.

CHAPTER 2

Solution Conformation of the Type I Collagen Alpha-2 Chain Telopeptides Studied by ^{1}H and ^{13}C NMR Spectroscopy*

Introduction

The collagens are a family of high molecular weight proteins produced and secreted by connective tissue cells such as fibroblasts, chondrocytes and osteoblasts, to perform various functions in the intercellular space. The most abundant form of collagen, genetic Type I, is found as relatively thick fibrils, approximately 20 to 200 nm or more in diameter, depending on tissue and age[1,2]. When organized into fibres and fibre bundles, these give mechanical strength to tissues such as dermis, tendon, bone and dentine. Fibrils closely resembling those seen *in vivo* precipitate from acidic solutions to pure monomeric Type I collagen on neutralization and warming[3,4]. Therefore the information necessary to direct this complex macromolecular assembly process is present in the sequence and higher orders of structure of the collagen monomer. Through most of its length the collagen molecule is a triple-helix of three polypeptide chains (two α-1 and one α-2 chain). The triple-helix is flanked by short non-triple-helical sequences (the "telopeptides"), 16 and 24 residues long in the case of the N- and C-terminal ends, respectively, of the α-1 chain, and 9 and 6 residues long for the N- and C-terminal ends of the α-2 chain of bovine type I collagen[5]. In spite of their relatively small size, the telopeptides

---

* A version of this chapter has been published. Liu, X., Scott, P. G., Otter, A. & Kotovych, G. (1990), *J. Biomol. Struct. Dyn.* **8**, 063-080.

seem to play critical roles in fibril assembly, since their removal by proteolysis seriously impairs this process[6,7]. These observations imply the existence of a set of specific, presumably conformation-dependent, non-covalent intermolecular associations between telopeptides on one molecule and the triple-helix of one or more adjacent collagen monomers within the growing fibril. The solution conformations of the telopeptides are therefore of considerable interest. We have recently described 2D-NMR studies on the α-1 chain N-[8] and C-telopeptides[9] which provide further insight into their roles in collagen fibrillogenesis. We present here the results of similar studies on both N- and C-telopeptides of the α-2 chain.

Materials and Methods

In the present study two telopeptides were studied, namely the 9-amino acid N-terminal α-2 chain telopeptide pGlu^1-Phe^2-Asp^3-Ala^4-Lys^5-Gly^6-Gly^7-Gly^8-Pro^9-NH_2 (pE^1-F^2-D^3-A^4-K^5-G^6-G^7-G^8-P^9-NH_2) and the C-terminal α-2 chain telopeptide N-acetyl-Ser^1-Gly^2-Gly^3-Tyr^4-Glu^5-Phe^6 (NAC-S^1-G^2-G^3-Y^4-E^5-F^6). Their structures are shown in Figure 2.1. Both peptides were synthesized commercially. The nonamer (IAF, Montreal, Quebec) was judged to be pure by HPLC and NMR. The hexamer (API, Edmonton, Alberta) was found to contain inorganic impurities which were removed by HPLC as described previously[9].

The NMR samples were prepared by dissolving the purified solid material in ca. 0.2 mL water (mixture of D_2O and H_2O). Then the pH was adjusted by adding small aliquots of diluted NaOH or HCl, followed by adding the rest of the solvent, which was 0.3 mL CD_3OH for most of the samples. To change the solvent at the same pH, the sample was freeze-dried and redissolved in the new solvent. Before the ROESY experiment the sample was degassed with argon.

The concentration of the nonamer was approximately 13 mM and of the hexamer 10 mM. The ^{1}H spectra were also run on the solutions which were diluted 10 times from the working samples under similar conditions, to confirm that there was no aggregation or other structural changes with concentration. The pH dependence of the ^{13}C resonances was carried out on a 28 mM sample. At pH 5.1 there was no dependence of the ^{13}C shifts on sample concentration.

All NMR experiments were carried out on Bruker AM400 and AM300 NMR spectrometers at ambient temperature (295K), except for the temperature study. The temperature dependence of the amide protons was calculated by comparing spectra recorded at 3 temperatures in the range from 273K to 293K with a maximum deviation of ± 0.3. Data collection was controlled by an Aspect 3000 computer using 1987 DISNMR software. For data processing the same computer and software were used, except that the Fourier transformations were carried out on the ASPECT 1000 computer with an array processor. The chemical shifts were measured relative to CD_3OH (Merck, Sharp & Dome, MD-67, 99.2% D) at 3.30 ppm (^{1}H) and 49.0 ppm (^{13}C), respectively. For the aqueous sample, a trace amount of CD_3OH was added as the chemical shift reference.

One-dimensional proton spectra were recorded with a sweep width of 20 ppm and 16K (zero-filled to 32K) data points in both water and water-methanol solvents. The presaturation of the solvent peak(s) was performed during a 3s relaxation delay at the lowest possible decoupler power to avoid saturation transfer and possible wipe-out of the exchangeable protons. For the sample in the CD_3OH/H_2O solutions, the frequency of the decoupler was changed every 0.3 s between two decoupler positions, one at the OH-signal and another at the signal of the undeuterated fraction of CD_3OH at 3.30 ppm. To account for the big difference in intensity between the two solvent resonances, the former signal

was saturated twice as long as the methyl group. The same basic procedure was followed for the aqueous sample, except that the decoupler was switched between three positions, namely on the water resonance and ± 5 Hz to either side of it. Finally, the resulting FID's were resolution enhanced by means of a Gaussian multiplication with a line broadening factor typically of -3 Hz.

The ^{13}C spectra were recorded in the mixed solvent using a DEPT (distortionless enhancement by polarization transfer) pulse sequence[10] for the protonated carbons. DEPT spectra were recorded with a sweep width of 160 ppm and 32K data points. The total relaxation delay was 1.8 s (including the acquisition time) and the total evolution delay for the polarization transfer 8.3 ms. The final proton detection pulse was 135°, creating positive signals for CH and CH_3 groups and negative signals for CH_2. Proton broadband decoupling was performed during the acquisition by means of a WALTZ[11] composite pulse technique. The FID was multiplied with an exponential function (line broadening factor 1.5) prior to Fourier transformation.

The parameters for the various two-dimensional NMR techniques are summarized in Table 2.1. The COSY[12], ROESY[13-17] and the ^{13}C-^{1}H correlation experiment using DEPT for polarization transfer[18] were all carried out in the phase-sensitive mode using time-proportional phase increments[19]. Solvent suppression was achieved in the COSY and ROESY experiments as described for the one-dimensional spectra. In addition, continuous irradiation was applied to the water signal at all times except during acquisition[20]. The phase constants were determined by using the second FID which was multiplied by the weighting function, Fourier transformed, and then phased. The phase constants thus determined were applied to all spectra in the t_2 domain. No phase correction was done in t_1 except for the ROESY experiments where a basic 90° phase correction was required. The resulting two-dimensional matrix was then usually

slightly rephased in order to obtain the best possible phase in the area of interest. The ROESY spectra showed some baseline distortions around intense diagonal peaks. Therefore, the spectra were subjected to a baseline correction routine[21], which fits a polynomial function of fifth degree to the baseline (ABS baseline correction), and the problem was eliminated almost completely in both dimensions.

Results

Proton NMR

The spectral analysis of the nonamer and the hexamer (Figure 2.1) is based on the results of the phase-sensitive COSY[19] and ROESY[13,14] experiments. The following discussion will detail the data for the nonamer but similar procedures were followed for the hexamer.

First of all, a pH titration was carried out in aqueous solutions (80% H_2O/20% D_2O) (Figure 2.2). It is clear that at a pH of 5.06 or higher the NH proton of D^3 has undergone a large shift because its β carboxyl group is now deprotonated. The solvent was then changed to 60% CD_3OH, 20% H_2O and 20% D_2O. This is a structure-inducing solvent mixture that we have used in the study of other peptides[8,9]. The contour plot of the COSY experiment in this solvent mixture is shown in Figure 2.3 (fingerprint region). The COSY experiment, combined with the ROESY (Figures 2.4a and 2.4b) experiment, allowed the complete assignment of all the protons of the nonamer. The spectrum was recorded with the transmitter approximately in the middle of the spectral range (ca. 5 ppm). $H_\alpha(i)$-NH(i+1) ROEs (Figure 2.4a) were obtained, allowing the sequential assignments. The biggest difficulty was in distinguishing the G^6 and G^7 proton resonances but this was resolved using the ROESY experiment. The results of the proton data are shown in Table 2.2, both at pH

5.06 as well as at a pH of 3.00. Since ROE cross-peaks show a non-negligible offset dependence[21], they often cannot be observed at the edge of a spectrum, irrespective of the internuclear distances involved. This is especially true for NH to NH interactions which are extremely important for the determination of secondary structures in peptides and proteins[22]. As such interactions were not observed with the transmitter at ca. 5 ppm, a ROESY experiment was performed with the transmitter set in the middle of the NH region. NH-NH ROE's, if present, would not have been observed because the D^3-A^4-K^5-G^6 NH-resonances are too close together and the cross peaks cannot be identified. The spin-lock conditions were achieved by a series of hard pulses of small flip angle[23] and the spin locking time was set to 100 ms at an average field strength of 4.0 kHz as was found suitable for peptides of similar size in previous studies. The spectra were basically free of any cross-peaks due to magnetization transfer between scalar-coupled spins (Hartmann-Hahn magnetization transfer)[14, 23].

The results of the spectral analysis of the proton data (Table 2.2) also includes the results of a temperature study of the exchangeable protons. The nonamer was recorded at four temperatures and the temperature coefficients were calculated. The temperature coefficient calculated for NH (G^6) of 3.2 ppb/K at a pH of 5.06 is indicative of shielding from the solvent, probably because of hydrogen bonding. When the pH is lowered to 3.0, unfortunately there is an overlap of the NH protons for G^6 and G^7 and a reliable temperature coefficient cannot be determined.

^{13}C NMR

The ^{13}C spectrum of the nonamer was first recorded for the protonated carbons by means of a DEPT[10] sequence applied to the 28.0 mM and 13.8 mM samples in water/methanol solutions. There was no concentration dependence

of the ^{13}C chemical shifts, within 0.1 ppm. Hence the resulting spectra are depicted in Figures 2.5a and 2.5b, as a function of pH at 28.0 mM. The doubling of the ^{13}C resonances is due to the fact that the backbone ^{13}C resonances see both NH and ND substitution and hence a two-bond isotope effect[24] due to the mixed solvent $CD_3OH/D_2O/H_2O$ (60/20/20). The magnitude of this isotope effect averages 0.11 ppm. The sample was then subjected to a $^{13}C/^{1}H$ two-dimensional phase-sensitive[18] correlation experiment. As can be seen in Figure 2.6 (a and b), all resonances can be assigned and the results are summarized in Table 2.3.

Because of the presence of two charged side-chain groups $Asp^3(D^3)$ and $Lys^5(K^5)$, a charge-charge interaction has to be considered. As can be seen from Figure 2.5, the $D_\alpha{}^3$ carbon undergoes a 0.91 ppm low-field shift due to deprotonation of its carboxyl group on changing the pH from 3.00 to 5.1. Over the same pH interval, $D_\beta{}^3$ undergoes a shift of +3.38 ppm. On the other hand, from pH 5.1 to pH 12.0, $K_\gamma{}^5$ shifts by +0.62 ppm, $K_\delta{}^5$ by +4.05 ppm and $K_\varepsilon{}^5$ by +1.41 ppm. These low-field shifts arise first from the deprotonation of the D^3 carboxyl group, followed by the deprotonation of the $NH_3{}^+$ of K^5. Hence, based on these data and in the absence of nuclear Overhauser effects between the two side-chains under consideration, it was difficult to prove or disprove the presence of a salt-bridge at pH values below the pK_a of $NH_3{}^+(K^5)$ and above the pK_a of $COO^-(D^3)$ (expected to be about 10.5 and 4.5, respectively).

The results for the hexamer will now be considered. The proton assignments were determined based on the COSY experiment, combined with the ROESY. The results, in $CD_3OH/D_2O/H_2O$ (60/20/20 volume percent) are presented in Table 2.4. The temperature dependence of the NH protons is also included. In the absence of a small NH temperature dependence, or small $J_{NH\text{-}CH\alpha}$ coupling constants or long-range ROE's, it was concluded that the hexamer

exists in an extended or random conformation. For completeness, the ^{13}C NMR data were also obtained. These are presented in Table 2.5.

Discussion

In principle, the NMR data (NH-H_α coupling constants, ROE distance constraints, amide temperature coefficients) provide the necessary information to estimate the torsion angles ϕ, ψ and ω[25] of the peptide backbone. Due to conformational averaging, which is always present in the peptides under consideration, the measured effects are the weighted average over two or more conformations making a numerically precise interpretation of, for example, the coupling constants impossible. Nevertheless, certain characteristics of the NMR parameters are diagnostic for the presence of secondary structure[22] namely, (i) small or large NH-H_α coupling constants of one or more residues, (ii) NH(i)-NH(i+1) nuclear Overhauser effect(s) due to short interproton distance(s), (iii) reduced temperature coefficient of one or more of the amide protons, and (iv) a low-field shift of some of the amide protons[22]. This low-field shift together with a reduced NH temperature coefficient is clearly indicative of the formation of a hydrogen bond, whereas the coupling constants and the ROEs reflect altered torsion angles ϕ and ψ. Depending on the nature of the non-extended secondary structure, and to what extent it coexists with different types of extended conformers, the magnitude of the effects might be small and therefore easily overlooked.

Conformation of the nonamer (N-telopeptide) in methanol/water

A review of the ^1H NMR data summarized in Table 2.2 reveals that this peptide is a representative of short linear peptides with some secondary structure. Most of the coupling constants and temperature coefficients as well as

the ROEs (Figure 2.7) are indicative of an extended conformation. However, the relatively small NH-H_α coupling constant measured in A^4 together with a reduced NH temperature coefficient of G^6 suggest that the peptide is either not completely extended or that a small proportion of the molecules exist in a folded conformation. Moreover, the low-field shift of NH(A^4) strongly suggests that C=O(D^3) takes part in a hydrogen bond. The most probable explanation of these observations would appear to be that the four residues D^3-A^4-K^5-G^6 of the nonamer form a Type I β-turn, stabilized by a C=O(D^3)-NH(G^6) hydrogen bond. The remaining residues are in an extended conformation. The magnitude of the effects discussed above and the observation of ROEs which are only consistent with an extended conformation, for example H_α(i)-NH(i+1) and the complementary H_α(i)-$H_{\delta\delta'}$(i+1) throughout the peptide chain, suggest that only a minor fraction of the molecules adopts the β-turn conformation whereas the majority is extended throughout. It is not possible to assess the relative populations of the molecules in the two conformations in a numerically precise way, but circular dichroism spectra (P.G. Scott, unpublished data) suggest that the proportion of molecules with defined secondary structure is very low - no more than 5-10% of the total. In theory, the NH-NH ROE cross peaks could be used to distinguish between a Type I and a Type II β-turn. As these could not be observed, we postulate a Type I β–turn because it is two to three times more common than the Type II, while the mirror images (Types I' and II') are very rare[26]. In addition, protein secondary structure predictions suggest a Type I turn for D^3-A^4-K^5-G^6 [26].

As can be expected in view of the G^8-P^9 subunit, the peptide bond angles ω are not uniformly trans (ω = 180°). Evidence for cis/trans isomerism[27,28] about the G^8-P^9 peptide bond is found in the ^1H and ^{13}C spectra. The cis-

isomer is present to the extent of about 10%. The chemical shifts related to the minor isomer are included in Tables 2.2 and 2.3.

Measurements were also carried out at low pH to see if the proportion of Type I β-turn changes. The proton data at low pH (3.00) (Table 2.2) prove to be inconclusive. Because of the overlap of the NH protons (G^6 and G^7), a reliable temperature dependence of these individual protons could not be measured. As well, NH(i)-NH(i+1) ROE's could not be detected at pH 3.00. Consequently, conclusions about the extent of secondary structure cannot be drawn.

Calculations using the ECEPP/2[29] (data not shown) indicate that an extended conformation for D^3-A^4-K^5-G^6, stabilized by a salt-bridge interaction between the ionized COO^- side chain of D^3 and the protonated NH_3^+ of K^5, would be of lower energy than the Type I β-turn structure. Some physical evidence in support of the existence of a salt bridge comes from the observation of the relatively high pH dependence of the chemical shift of the NH proton of A^4 - the intervening residue (Table 2.2).

Conformation of the hexamer (C-telopeptide) in CD_3OH/water

The proton data in Table 2.4 show neither small $J_{NH\text{-}CH\alpha}$ coupling constants nor very low NH temperature coefficients. Consequently we conclude that this peptide is extended throughout.

Implications for collagen telopeptide-triple helix interactions

In spite of their relatively small size compared to the telopeptides of the α-1 chain, the α-2 chain telopeptides may nevertheless contribute to collagen fibrillogenesis. Linear growth of fibrils is driven by interactions between molecules displaced by 4D (936 helical residues) and lateral growth by 1D (234 helical residues) interactions. Experiments on proteolytically-modified collagen

strongly suggest that the N-telopeptides are primarily involved in linear growth and the C-telopeptides in lateral growth[30,31]. It can be see in Figure 2.8 that there is a greater potential for both charge-charge and hydrophobic interactions between the α-2 chain N-telopeptide and the adjacent segment of triple-helix displaced by 4D, rather than 1D. The converse is true for the C-telopeptide. In both cases an extended conformation would seem to result in better positioning of the telopeptide side-chains for interaction than would a folded structure involving one or more β-turns. Nevertheless, such structures have been predicted by others from the amino acid sequence[32]. The physical evidence from the NMR studies described here for an extended conformation of the α-2 chain telopeptides does not support the predictions of probability-based algorithms but is in accord with the above observations on potential telopeptide-helix interactions.

References

1. Parry, D. A. D., Craig, A. S. & Barnes, G. R. G. (1978), *Proc. Royal Soc. (London)* **203**, 305-321.
2. Parry, D. A. D. (1984), *Ultrastructure of the Connective Tissue Matrix*, Ruggeri, A. & Motta, P. M. Eds., Martinus-Nijhoff, Boston, 34-64.
3. Gross, J., Highberger, J. H. & Schmitt, F. O. (1955), *Proc. Natl. Acad. Sci. U.S.A.* **41**, 1-7.
4. Jackson, D. S. & Fessler, J. H. (1955), *Nature (London)* **176**, 69-70.
5. Glanville, R. W. & Kühn, K. (1979), *Fibrous Proteins: Scientific, Industrial and Medical Aspects*, Parry,D. A. D. & Cramer, L. K. Eds., Academic Press, London, **1**, 133-150.
6. Rubin, A. L., Pfahl, D., Speakman, P. T., Davison, P. F. & Schmitt, F. O. (1963), *Science (Washington, DC)* **139**, 37-39.
7. Drake, M. P., Davison, P. F., Bump, S. & Schmitt, F. O. (1966), *Biochemistry* **5**, 301-310.
8. Otter, A., Kotovych, G. & Scott, P. G. (1989), *Biochemistry* **28**, 8003-8010.
9. Otter, A., Kotovych, G. & Scott, P. G. (1988), *Biochemistry* **27**, 3560-3567.
10. Doddrell, D. M., Pegg, D. T. & Bendall, M. R. (1982), *J. Magn. Reson.* **48**, 323-327.
11. Shaka, A. J., Keeler, J. & Freeman, R. (1983), *J. Magn. Reson.* **53**, 313-340.
12. Aue, W. R., Bartholdi, E. & Ernst, R. R. (1976), *J. Chem. Phys.* **64**, 2229-2246.
13. Bothner-By, A. A., Stephens, R. L., Lee, J., Warren, C. D. & Jeanloz, R. W. (1984), *J. Am. Chem. Soc.* **106**, 811-813.
14. Bax, A. & Davis, D. G. (1985), *J. Magn. Reson.* **63**, 207-213.

15. Brown, L. R. & Farmer II, B. T. (1989), *Methods in Enzymology* **176**, 199-216.

16. Otter, A. & Kotovych, G. (1988), *Can. J. Chem.* **66**, 1814-1820.

17. Fejzo, J., Zolnai, Z., Macura, S. & Markley, J. L. (1989), *J. Magn. Reson.* **82**, 518-528.

18. Bendall, M. R. & Pegg, D. T. (1983), *J. Magn. Reson.* **53**, 144-148.

19. Marion, D. & Wüthrich, K. (1983), *Biochem. Biophys. Res. Comm.* **113**, 967-974.

20. Wider, G., Hosur, R. V. & Wüthrich, K. (1983), *J. Magn. Reson.* **52**, 130-135.

21. Griesinger, Ch. & Ernst, R. R. (1987), *J. Magn. Reson.* **75**, 261-271.

22. Wüthrich, K. (1986), *NMR of Proteins and Nucleic Acids*, Wiley-Interscience, New York.

23. Kessler, H., Griesinger, Ch., Kerssebaum, R., Wagner, G. & Ernst, R. R. (1987), *J. Am. Chem. Soc.* **109**, 607-609.

24. Otter, A., Liu, X. & Kotovych, G. (1990), *J. Magn. Reson.* **86**, 657-662.

25. IUPAC-IUB Commission on Biochemical Nomenclature (1970), *Biochemistry* **9**, 3471-3479.

26. Wilmot, C. M. & Thornton, J. M. J. (1988), *Mol. Biol.* **203**, 221-232.

27. Deslauriers, R. & Smith, I. C. P. (1976), "*Topics in Carbon-13 NMR Spectroscopy*, Levy, G. C., Eds., Wiley-Interscience, New York, 2-66.

28. Otter, A. & Kotovych, G. (1987), *J. Magn. Reson.* **74**, 293-307.

29. Nemethy, G., Pottle, M. S. & Scheraga, H. A. (1983), *J. Phys. Chem.* **87**, 1883-1887.

30. Helseth, Jr., D. L. & Veis, A. (1981), *J. Biol. Chem.* **256**, 7118-7128.

31. Capaldi, M. J. & Chapman, J. A. (1982), *Biopolymers* **21**, 2291-2313.

32. Jones, E. Y. & Miller, A. (1987), *Biopolymers* **26**, 463-480.

33. Chapman, J. A. & Hulmes, D. J. S. (1984), *Ultrastructure of the Connective Tissue Matrix*, Ruggeri, A. & Motta, P. M., Eds., Martin-Nijhoff, Boston, 1-33.

34. Traub, W. & Fietzek, P. (1976), *FEBS Letters*. **68**, 245-249.

Table 2.1. Summary of Experimental Parameters Used in the Two-Dimensional NMR Experiments

Parameters	COSYPH	ROESYPH	XHDEPTPH
Sweep width in F2 (Hz)	3600	3600	5000
Sweep width in F1 (Hz)	1800	1800	1400
Matrix size (F1 x F2) before zero-filling	512 x 4K	512 x 1K	64 x 4K
Matrix size (F1 x F2) after zero-filling	1K x 8K	1K x 2K	256 x 4K
Evolution time			
Initial value (us)	1	1	1
Increment (us)	139	139	178
Number of scans (dummy scans)	56(2)	56(2)	640(2)
Acquisition time (s)	0.57	0.57	0.41
Relaxation delay (s)	1.8	1.8	1.4
Other delays (ms)		100[b]	4.2[c]
Window functions for 2D FT(F1/F2)	S/S[d]	S/S	E/E[d]
Shifts of window function in fractions of π(F1/F2)	6/8	3/4	3.0/1.5[e]

(a) For abbreviation of techniques, see text.

(b) Spin locking time at an average field strength of 4.0 KHz.

(c) $(2J_{^{13}C - ^{1}H})^{-1}$

(d) S stands for a sine bell, E for an exponential function.

(e) Line broading factors in F1 and F2.

Table 2.2. Proton Chemical Shifts, Coupling Constants and NH Temperature Dependence of pE-F-D-A-K-G-G-G-P-NH_2 [a]

	NH	α	β	γ	Others	J (Hz)	ΔT (ppb/K)
pE[1]	7.86	4.18	1.72,2.38	2.13,2.26			
F[2]	8.23	4.67	2.93,3.19			8.6	6.9
	(8.28)					(7.0)	(8.1)
D[3]	8.34	4.58	2.67			8.6	5.2
	(8.43)					(7.7)	(6.5)
A[4]	8.37	4.21	1.41			5.7	7.2
	(8.17)					(5.5)	(6.4)[c]
K[5]	8.31	4.27	1.86	1.45	δ:1.68	7.4	5.3
					ε:2.97		
	(8.17)					(7.8)	(6.4)[c]
G[6]	8.23	3.92				6.0/6.0	3.2
	(8.26)					(5.5/5.5)	(5.3) [d]
G[7]	8.31	3.96				6.0/6.0	5.9
	(8.26)					(5.5/5.5)	(5.3) [d]
G[8b]	8.10	4.07				6.0/6.0	5.2
	(8.14)					(5.8/5.8)	(6.2)
P[9b]		4.36	1.96,2.22	2.00	δ:3.54		
					δ:3.64		
NH_2[b]	6.97, 7.68						
	(6,99,7.70)						

(a) The nonamer was studied at 13.8 mM in $CD_3OH/H_2O/D_2O$ (60/20/20). The pH was 5.06, whereas the data in brackets are at pH 3.0. The shifts are in ppm with respect to undeuterated CD_3 of CD_3OH at 3.30 ppm

(b) The chemical shifts for the cis-isomer are: G^8NH, 8.13; G^8CH_{α}, 3.72; P^9CH_{α}; 4.48; NH_2 7.88, in ppm.

(c) Two sets of data are not distinguishable because of spectral overlap.

(d) Two sets of data are not distinguishable because of spectral overlap.

Table 2.3. ^{13}C Chemical Shifts of pE-F-D-A-K-G-G-G-P-NH_2 as a Function of pH[a]

	pH	α	β	γ	δ	ε	ζ
pE[1]	3.00	57.78	26.21	29.94			
	5.13	57.90	26.21	29.92			
	12.00	57.91	26.17	29.93			
F[2]	3.00	56.07	37.82		130.00	129.50	127.98
	5.13	55.83	38.02		130.04	129.48	127.94
	12.00	55.84	37.93		130.05	127.94	128.01
D[3]	3.00	51.08[b]	36.43				
	5.13	51.99	39.81				
	12.00	52.35	39.74				
A[4]	3.00	51.08[b]	17.38				
	5.13	51.44	17.15				
	12.00	51.32	17.35				
K[5]	3.00	54.68	31.22	23.10	27.43	40.33	
	5.13	54.98	30.87	23.11	27.34	40.21	
	12.00	55.20	32.22	23.73	31.39	41.62	
G[6]	3.00	43.52					
	5.13	43.74					
	12.00	43.75					
G[7]	3.00	43.28					
	5.13	43.32					
	12.00	43.34					
G[8]	3.00	42.62(42.40)[c]					

	5.13	42.86(42.37)			
	12.00	42.68(42.38)			
P^9	3.00	61.37	30.56	25.30	47.69
	5.13	61.42	30.54	25.29	47.72
	12.00	61.42	30.55	25.30	47.77
P^9(cis)	3.00	60.91	33.03	23.00	
	5.13	60.92	33.00	23.00	
	12.00	60.95	33.00	23.00	

(a) The samples were 28.0 mM in $CD_3OH/D_2O/H_2O$ (60/20/20). Experiments carried out at 13.8 mM indicate that the shifts are the same, within experimental error (0.1 ppm). Because the backbone resonances experience an isotope effect due to the presence of both NH and ND, the chemical shifts of the more intense peaks (NH substitution two bonds away) are presented.

(b) The two signals cannot be distinguished because of spectral overlap.

(c) The data in brackets are for the cis isomer.

Table 2.4. Proton Chemical Shifts, Coupling Constants and NH Temperature Dependence of NAc-S-G-G-Y-E-F[a]

	NH	α	β	γ	J (Hz)	ΔT (ppb/K)
S1[b]	8.34	4.39	3.79,3.88		7.0	8.4
G2	8.59	3.92			6.5/6.5	7.5
G3	8.26	3.84			6.0/6.0	5.1
Y4	7.94	4.52	2.80,2.99		8.0	6.4
E5	8.39	4.22	1.29,1.82	4.03	7.8	6.3
F6	7.47	4.40	2.94,3.16		8.0	4.4

(a) The sample was 10.0 mM, pH 5.5 in $CD_3OH/D_2O/H_2O$ (60/20/20).

(b) The chemical shift of N-acetyl is 2.05 ppm.

Table 2.5. ^{13}C Chemical Shifts of NAc-S-G-G-Y-E-F[a]

	α	β	γ	δ	ε	ζ
S^{1}[b]	57.30	62.51				
G^{2}	43.80					
G^{3}	43.28					
Y^{4}	56.20	37.26		131.37	116.31	
E^{5}	55.12	28.94	34.68			
F^{6}	57.04	38.75		130.42	129.25	127.46

(a) The solvent is $CD_3OH/D_2O/H_2O$ (60/18/22 volume %). The sample concentration was ~ 10 mM, pH 5.3. Because the backbone resonances experience an isotope effect due to the presence of both NH and ND, the chemical shifts of the more intense peaks are presented (NH substitution two bonds away).

(b) The chemical shift of N-acetyl is 22.74 ppm.

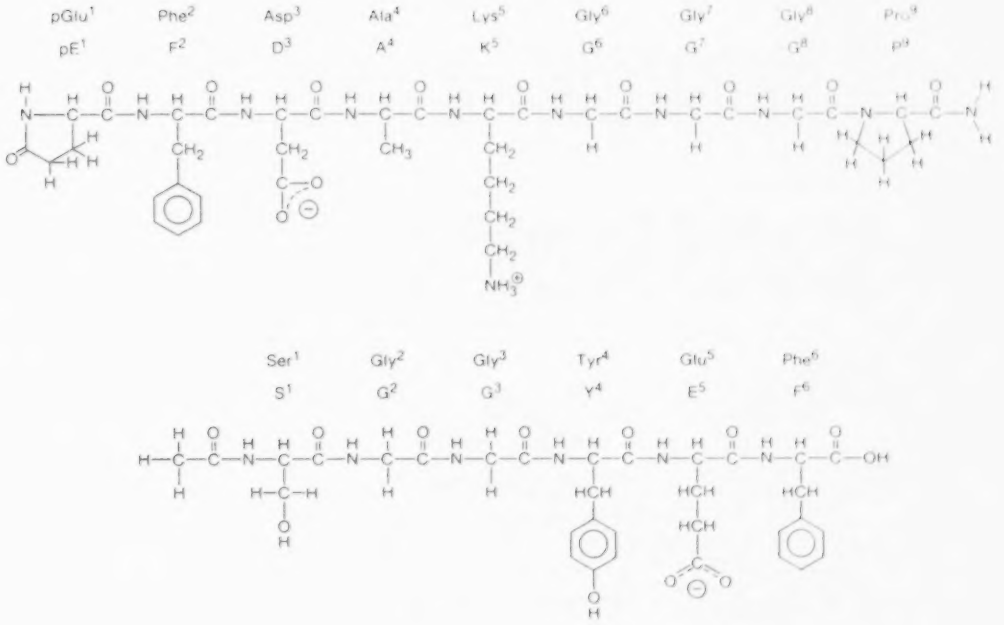


Figure 2.1 The structures of the nonamer and the hexamer, namely the N-terminal and C-terminal Type I, α-2 chain telopeptides.

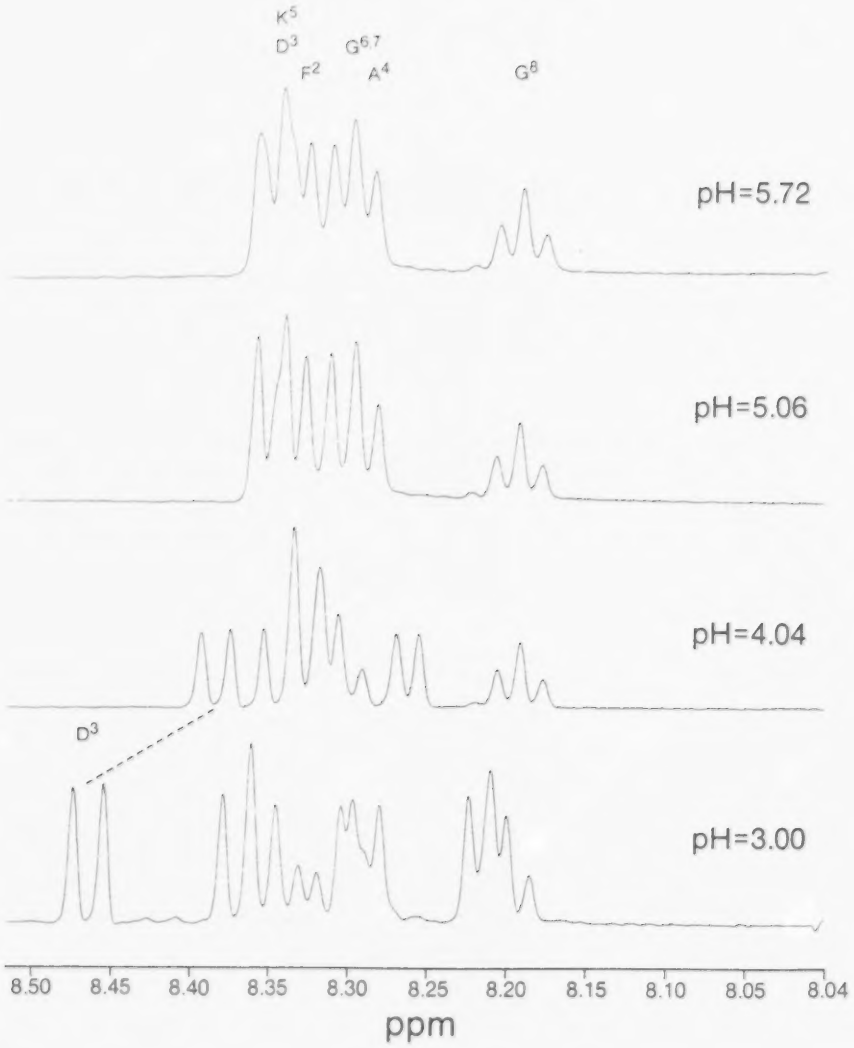


Figure 2.2 The pH titration of the nonamer in aqueous solution. Only the NH protons are shown. The solution concentration was 13 mM.

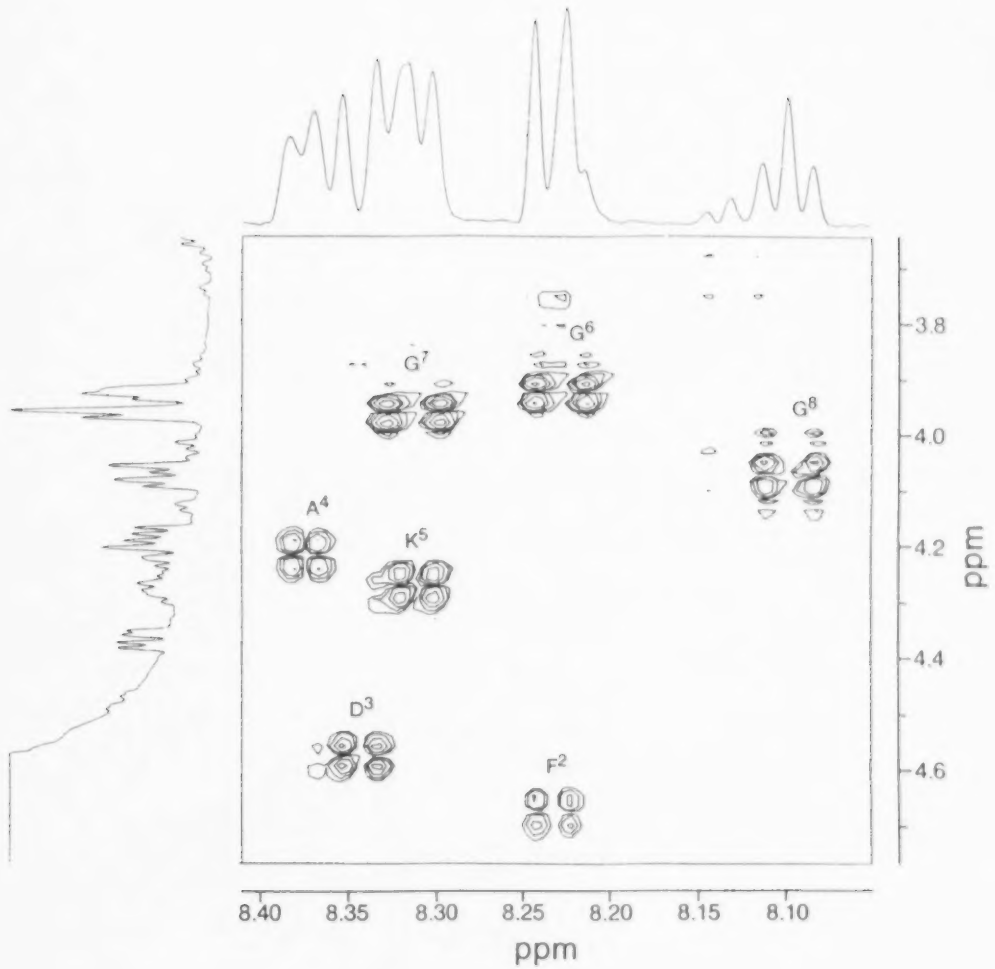


Figure 2.3 The COSY spectrum of the nonamer. The fingerprint region is presented. The solvent is $CD_3OH/H_2O/D_2O$ (60/20/20 volume percent). The connectivities are labelled in the figure. The pH was 5.06. The chemical shifts are listed in Table 2.2.

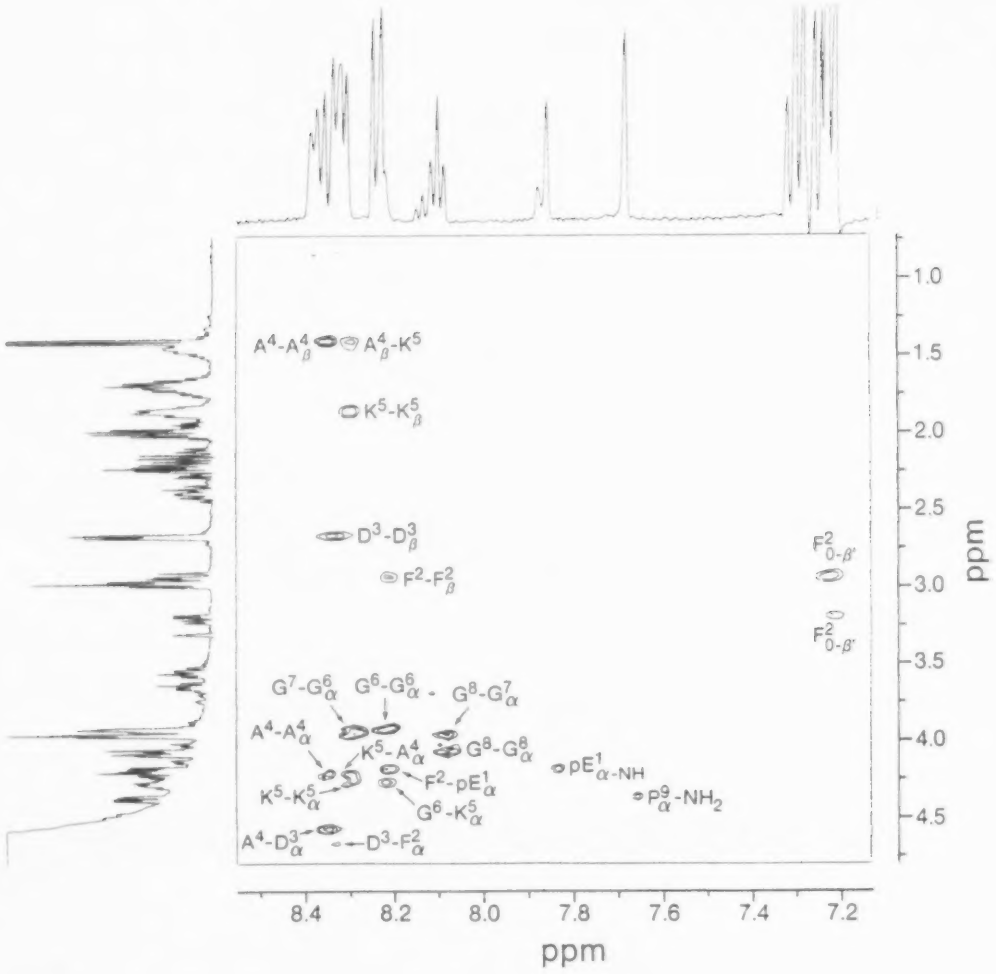


Figure 2.4a The low-field region of the ROESY spectrum of the nonamer under the same conditions as Figure 2.3. The contour levels are of opposite sign with respect to the diagonal peaks.

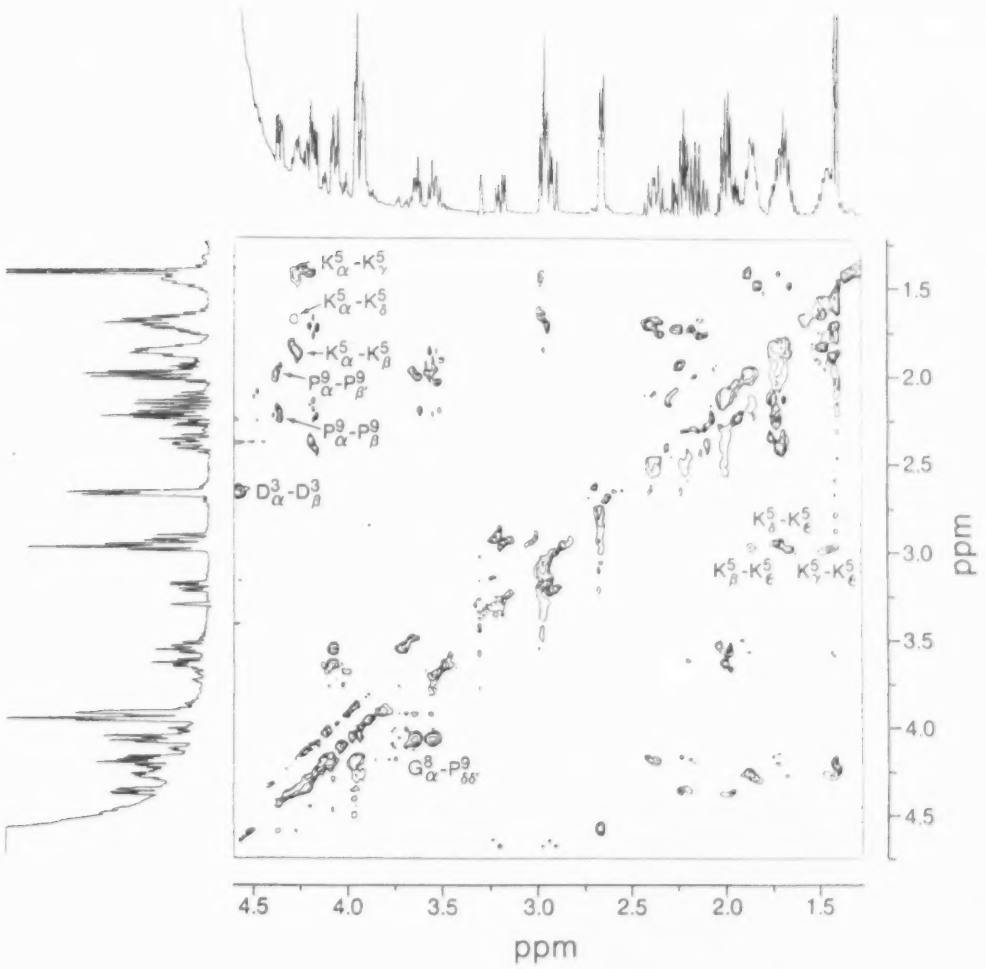


Figure 2.4b The high-field region of the ROESY spectrum of the nonamer under the same conditions as Figure 2.3. The contour levels are of opposite sign with respect to the diagonal peaks.

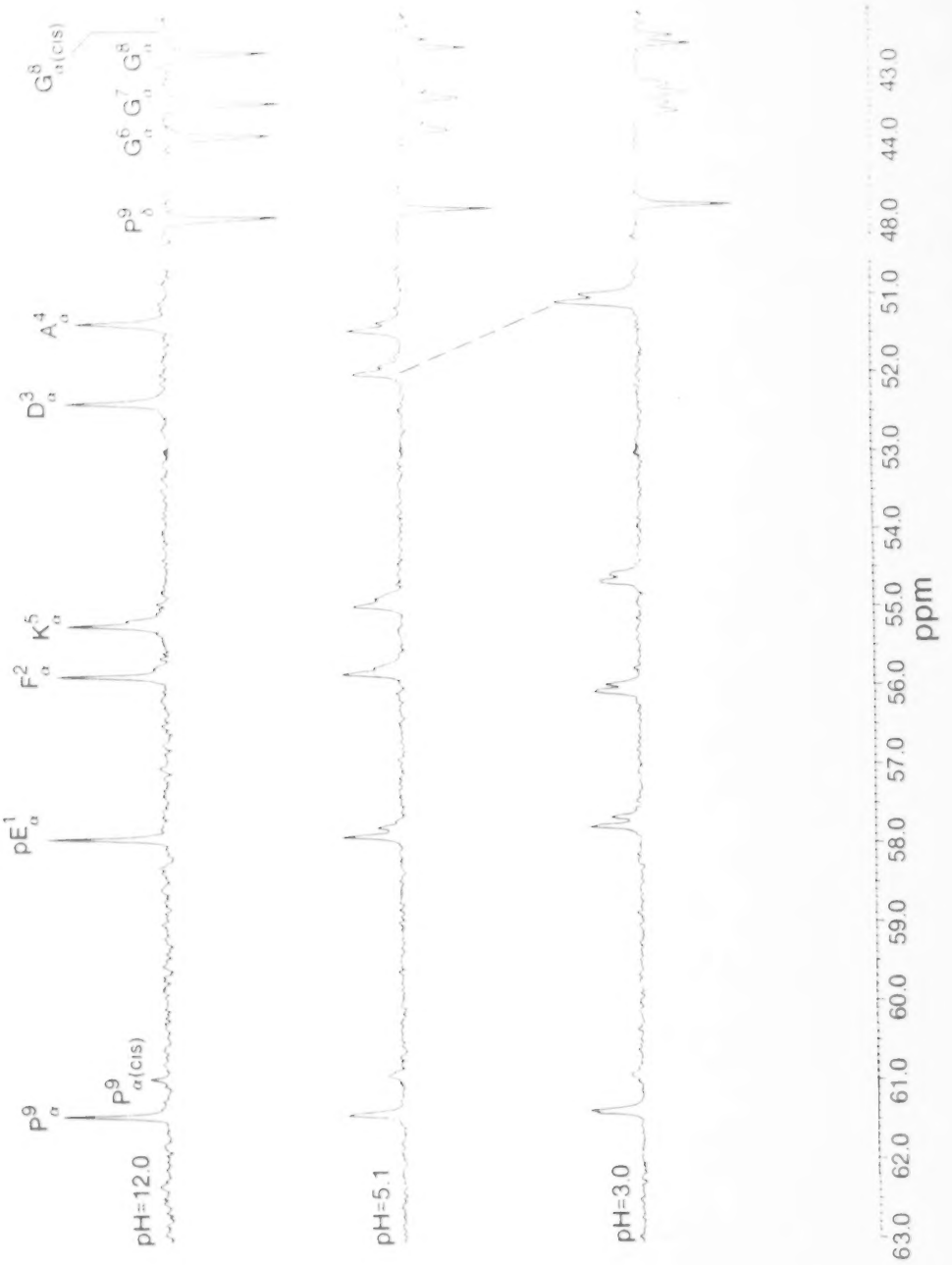


Figure 2.5a The low-field ^{13}C spectrum of the nonamer as a function of pH. CH and CH_3 carbons are phased positive whereas CH_2 carbons are phased negative. The data are presented in Table 2.3.

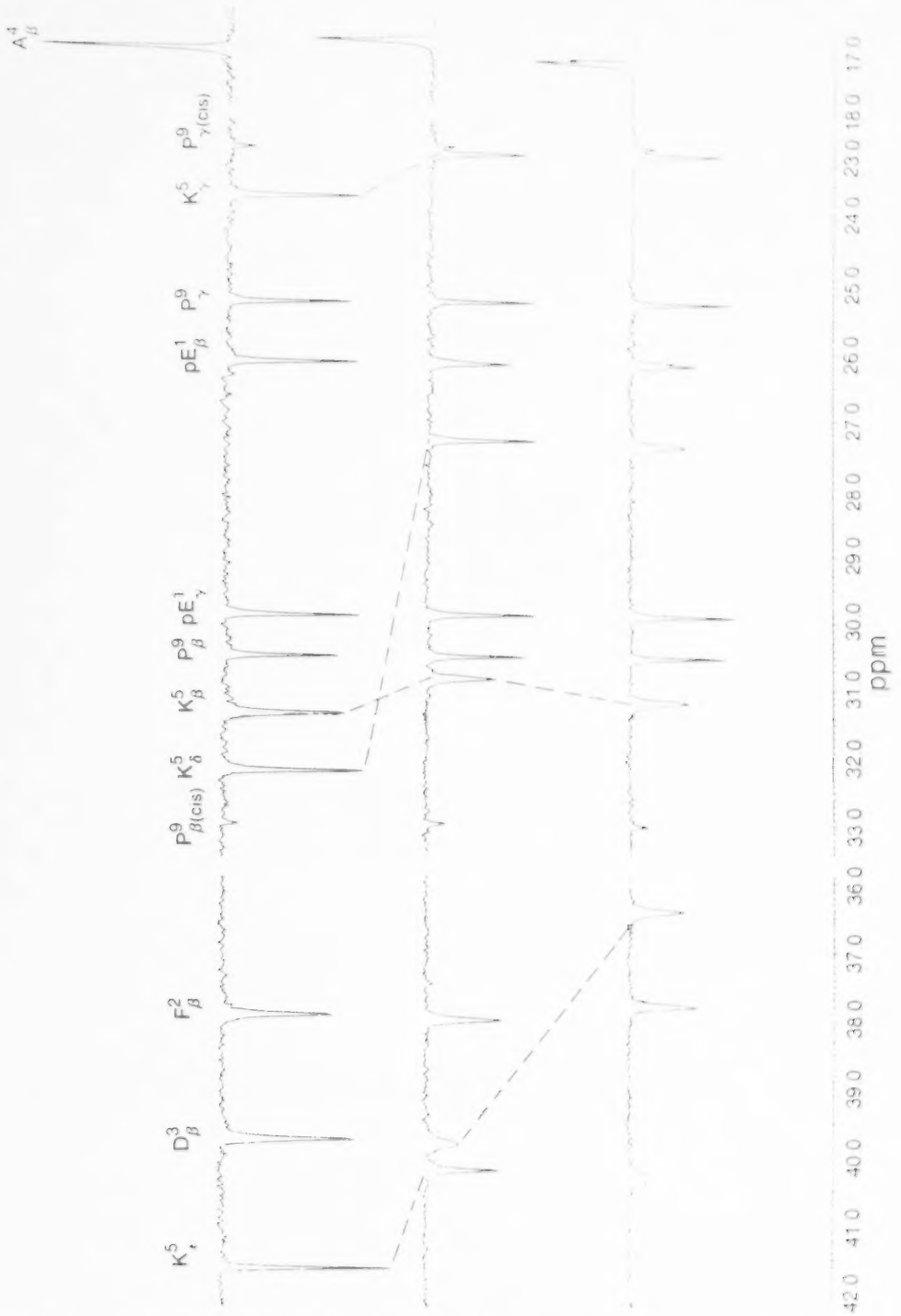


Figure 2.5b The high-field ^{13}C spectrum of the nonamer as a function of pH. CH and CH_3 carbons are phased positive whereas CH_2 carbons are phased negative. The data are presented in Table 2.3.

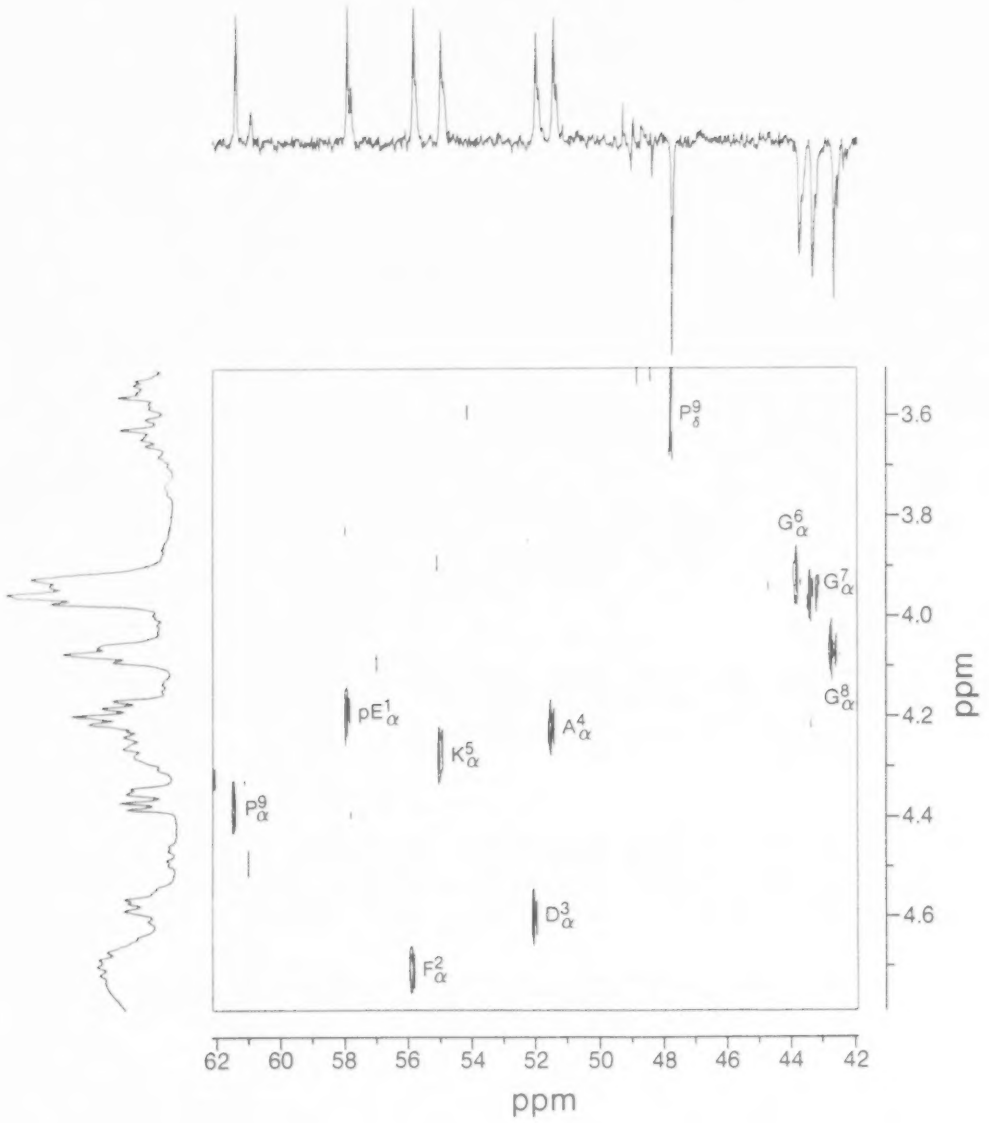


Figure 2.6a The low-field region of the ^1H-^{13}C correlation spectrum of the nonamer in $CD_3OH/D_2O/H_2O$ (60/20/20 volume percent) at 28.0 mM and pH 5.13. The chemical shifts are presented in Table 2.3.

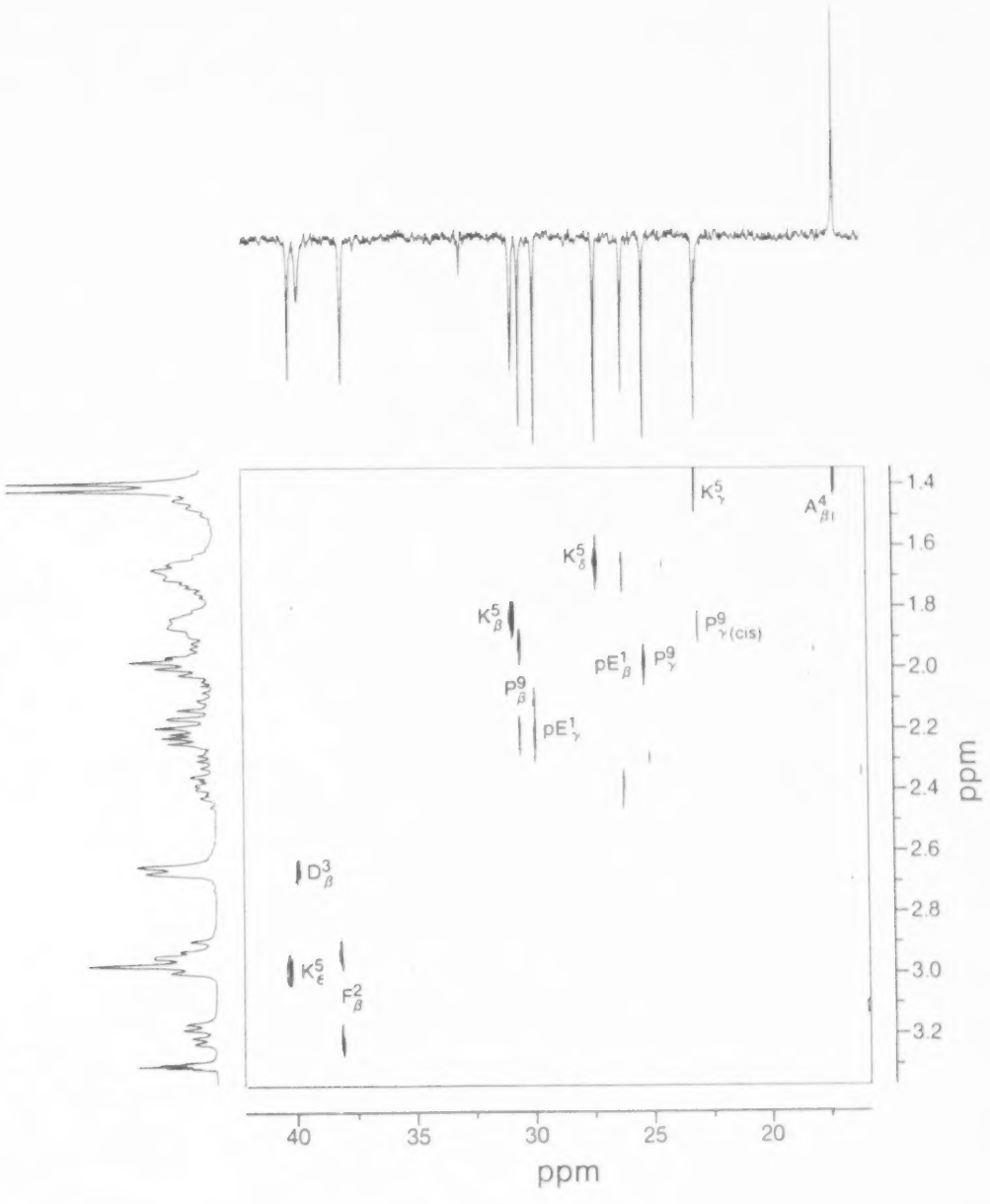


Figure 2.6b The high-field region of the ^1H-^{13}C correlation spectrum of the nonamer in $CD_3OH/D_2O/H_2O$ (60/20/20 volume percent) at 28.0 mM and pH 5.13.spectrum. The chemical shifts are presented in Table 2.3.

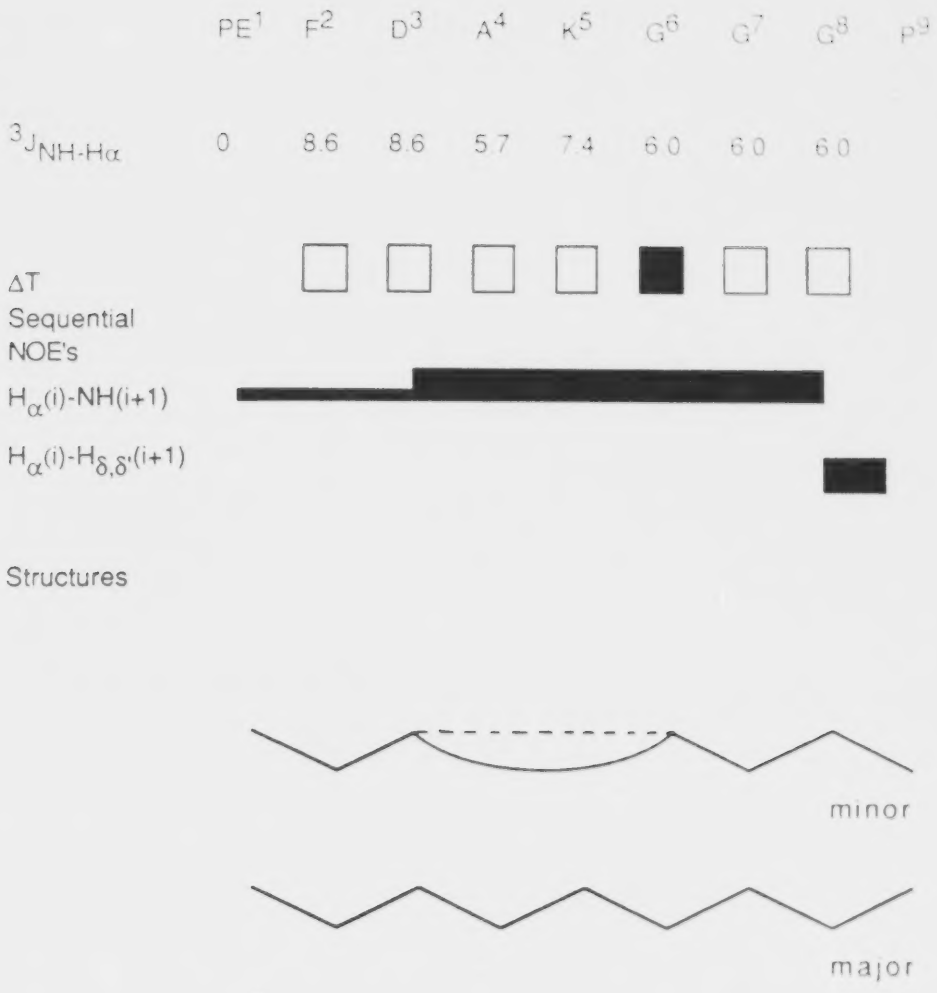


Figure 2.7 The results of the proton NMR data in CD_3OH/water, including the ROE data, of the nonamer, together with the two proposed structures, a type I β-turn and the extended structure. The $^3J_{NH\text{-}H\alpha}$ coupling constants are in Hz. The solid square indicates a low temperature dependence of the NH protons (3.2 ppb/K), whereas the open squares indicate a high temperature dependence (>5.2 ppb/K).

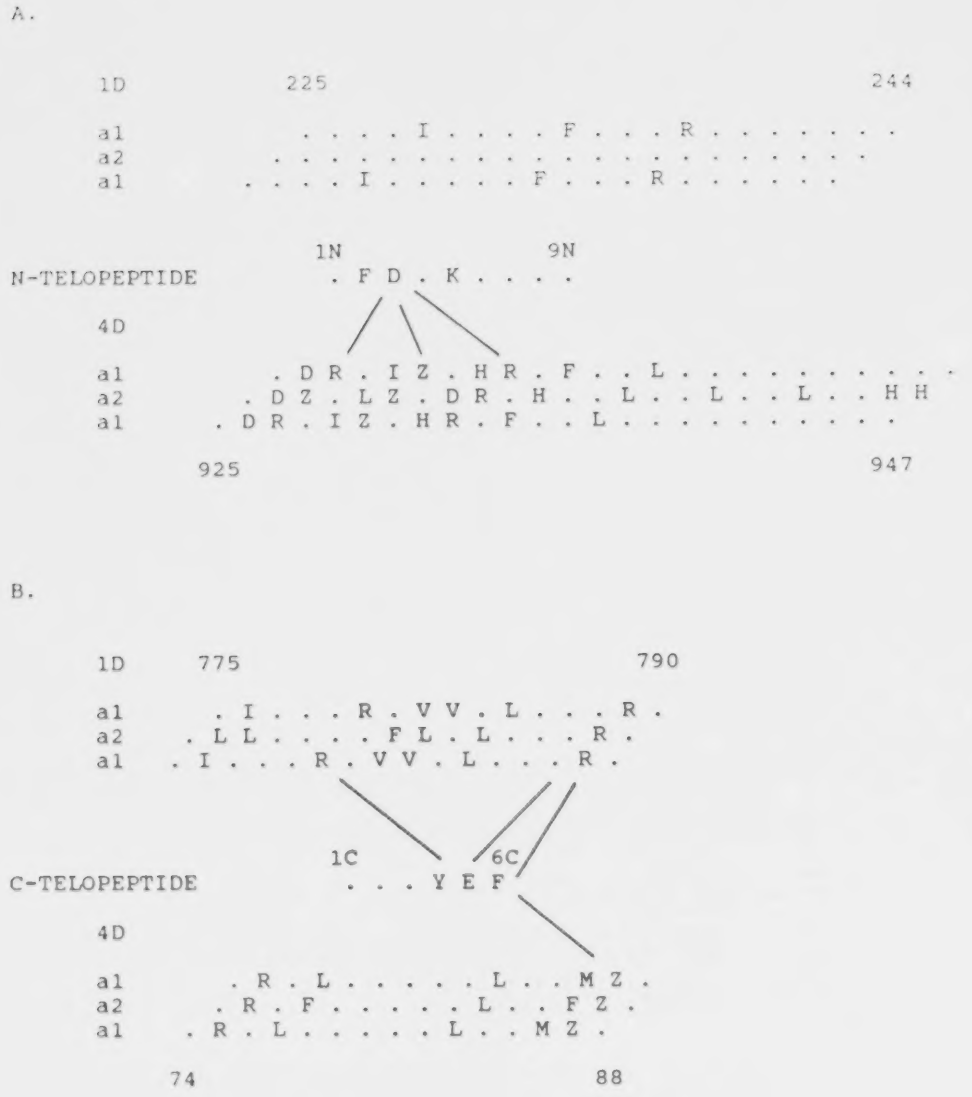

A.
1D
225
244
a1 . . . . I . . . . F . . . R . . . . . .
a2 . . . . . .
a1 . . . . I . . . . F . . . R . . . . . .
1N
9N
N-TELOPEPTIDE . F D . K . . . .
4D
a1 . D R . I Z . H R . F . . L . . . . . .
a2 . D Z . L Z . D R . H . . L . . L . . L . . H H
a1 . D R . I Z . H R . F . . L . . . . . .
925
947
B.
1D
775
790
a1 . I . . . R . V V . L . . . R .
a2 . L L . . . . F L . L . . . R .
a1 . I . . . R . V V . L . . . R .
1C
6C
C-TELOPEPTIDE . . . Y E F
4D
a1 . R . L . . . . . L . . M Z .
a2 . R . F . . . . . L . . F Z .
a1 . R . L . . . . . L . . M Z .
74
88

Figure 2.8 Alignment of N- (A) and C-terminal (B) α-2 chain telopeptides with sequences[33] of triple-helices displaced axially by 1D (234 residues) and 4D (936 residues). Hydroxylysine is represented by Z. Only potentially interacting (charged and large hydrophobic) residues are shown. The α-2 chain is assumed to be in the middle (reference) position in the triple-helix[34]. The solid lines show potential favourable charge-charge interactions.

CHAPTER 3

Conformational Analysis of the Type II and Type III Collagen α-1 Chain C-Telopeptides by ^{1}H NMR and Circular Dichroism Spectroscopy

Introduction

Collagen is the protein that gives mechanical strength to cartilages, bones, skin, tendons, ligaments and teeth. The short N- and C-terminal non-triple-helical extensions ("telopeptides") present on each of the three α-chains of the fibrillar collagen monomers play key roles in the assembly of the collagen fibrils which give connective tissues their mechanical strength. Proteolytic removal of telopeptides (which range from 6 to 27 amino acid residues long, depending on the particular α-chain and collagen type) either abolishes the ability of the collagen monomers to assemble or leads to fibrils with abnormal morphology. From these results it appears that the N-telopeptides of type I collagen direct linear growth of the fibrils[1], while C-telopeptides promote lateral growth[2,3]. Fibrillogenesis is believed to involve their interaction in the growing fibril with a specific region on the triple-helix of an adjacent monomer. Protein-protein interactions are by nature conformation-dependent and we have recently completed a series of NMR studies on the conformations in solution of all the N-telopeptides of type I, II and III collagens, which are most abundant collagen types in the body, as well as the type I C-telopeptide[4-8]. The N-telopeptide of type I collagen forms a significant proportion of β-turn in an environment of reduced dielectric constant (60% methanol) which is thought to mimic conditions within the growing fibril[6]. Under the same conditions the type II collagen N-

telopeptide forms a fused β/γ turn and the type III collagen N-telopeptide forms a "nascent helix"[8].

Type I collagen is found in bones, teeth, skin, tendons and ligaments, whereas type II collagen is found in adults only in hyaline cartilage, where it is the dominant form, and in vitreous humour. *In vivo*, it is present as thin fibrils which are critical components in the physical integrity of cartilage. The N- and C-telopeptides of human type II collagen are 19 and 27 residues long, respectively[9]. Type III is the second most abundant form of collagen in soft tissues other than cartilage. It is prevalent in tissues where distensibility is a major requirement for function, such as blood vessel walls, lung parenchyma, intestine and uterine wall[10]. *In vitro*, type III collagen has been shown to form thinner fibrils than type I and to reduce the diameter of copolymeric fibrils containing predominantly type I.[11] *In vivo*, it appears that the fibrils in many tissues contain both types I and III[12]. The type III collagen monomer consists of three identical α-1 chains. Sequence data are available for the bovine type III N-telopeptide, which is 14 residues long[13] and the human[14] and chick[15] C-telopeptides which are 25 and 23 residues long, respectively.

While there have been attempts to predict secondary structures of collagen telopeptides[16] there exist few supporting physical data. To complement our previous studies, the present paper presents the results of our two-dimensional NMR and circular dichroism studies of the type II and type III α-1 chain collagen C-telopeptides.

Materials and Methods

Materials and sample preparation

Both peptides, the bovine type II NAc-GPGIDMSAFAGLGPRE KGPDPLQYMRA and the chick type III NAc-GGGVASLGAGEKGPVGYGYEYR,

were synthesized on a Biolynx™ 4175 manual peptide synthesizer using Fmoc amino acid derivatives (Pharmacia Canada Inc., Baie d'Urfe, Quebec). The raw materials were purified by reversed-phase HPLC on a Vydac TP218 column (The Separations Group, Hesperia, CA). The resulting white products were judged to be pure, based on their amino acid compositions and NMR spectra.

The peptide content (71.8 %) of the type II collagen C-telopeptide was determined by spectroscopic analysis for the tyrosine content of an aqueous solution containing 0.300 mg peptide preparation per mL at pH 5.4. The near-UV spectrum of the solution in a 1 cm pathlength cell was recorded on a Beckman Model 25 Spectrophotometer, the concentration of tyrosine being calculated from the absorbance at 275 nm and the molar extinction coefficient, 1420 L/mol cm, of N-acetyl tyrosine methyl ester.[17]

A 4.2 mg sample of the type II C-telopeptide was dissolved in 140 μL of H_2O. The pH was measured in the 5 mm NMR tube with a Cole-Parmer C-5990 electrode and adjusted to 5.4 by adding small aliquots of 0.2 M NaOH. The rest of the solvent, 560 μL of CD_3OH (Merck, Sharp and Dome, MD-67, 99.9% D), was then added and degassing of the solution with argon was performed prior to sealing. The final concentration was ca 2.1 mM in a mixture of 80% CD_3OH and 20% H_2O.

A 4.9 mg sample of the type III C-telopeptide was dissolved in 140 μL of H_2O. The solution was filtered and transferred to a 5 mm NMR tube and the pH was adjusted to 5.5 as described above. The rest of the solvent, 560 μL of CD_3OH, was then added and the solution was degassed with argon prior to sealing. The final concentration was 3.0 mM in a mixture of 80% CD_3OH and 20% H_2O.

Circular dichroism spectroscopy

The CD spectra shown in Figure 3.1 were recorded at 10 ± 0.05 °C on a Cary model 60 spectropolarimeter with a model 6001 circular dichroism attachment with thermostated cell holder. The instrument had been modified to eliminate possible artifactual signals on passing through intense absorption bands. Slits were programmed for 1.5 nm bandwidth at each wavelength. The instrument was calibrated with dextro-10-camphor-sulfonic acid using the procedure suggested by Chen and Yang.[18] The path length, 0.05 to 2 cm depending upon wavelength range, was dictated by the absorbance of the solution. Mean residue ellipticities, $[\Theta]_{mrw}$(deg cm^2 $dmol^{-1}$), were calculated in the usual fashion using the mean residue weight of 106, calculated from the molecular weight of the peptide. Each spectrum is the average of two scans.

NMR spectroscopy

All NMR experiments were carried out on a Varian UNITY500 spectrometer using a proton-selective 5 mm probe with a 90° proton pulse length of 9.0 μs at a transmitter power level of 60 dB. The 90° pulse length was checked at the beginning of every NMR session and the temperature was always controlled to ±0.1°C. For data collection and processing, VNMR software (version 3.2 and 4.1) was used on Sun Sparc 4/330 and Sparc 4/470 workstations, respectively. 2D spectra were acquired non-spinning, whereas for one-dimensional experiments the spinner was turned on. All chemical shifts are measured relative to the undeuterated fraction of the methyl group of CD_3OH at 3.30 ppm. The temperature dependent coefficients of the amide protons were calculated by analyzing the chemical shifts at 0.0, 5.0, 10.0, 15.0, 20.0 and 25.0°C (type II) and 0.0, 10.0, and 20.0°C (type III) by means of a linear

regression. Unless otherwise indicated, all data are reported at 0.0°C for the type II N-telopeptide and 10.0°C in case of the type III N-telopeptide.

All two-dimensional NMR experiments[19] were carried out in the phase-sensitive mode by using the hypercomplex technique (States-Haberkorn-Ruben method.)[20] The decoupler and transmitter offsets were always set equal, a sweep width of 5000 Hz was used in both dimensions and typically 4K data points in t2 and 512 experiments in t1 (zero-filled to 1K) were acquired. The only significant exceptions from this scheme were the TOCSY[21] experiments where 800 t1-increments were acquired and an 8K by 2K data matrix was analyzed. All two-dimensional experiments were preceded by 8 dummy scans but none were applied between individual t1-increments. The preacquisition delay "alfa"[22] was carefully and individually adjusted for every 2D experiment. First point distortions caused by analog filters[23] were compensated by reducing the first point of each FID and of each t1-interferogram by empirically determined values.[22] No further correction routines were necessary.

The TOCSY experiments were carried out by using the basic pulse sequence proposed by Bax.[24] Water suppression was achieved by solvent presaturation during the relaxation period. A 2 ms trim pulse preceded the 80 ms MLEV-17 spin-lock (field strength 7.0 kHz). Connectivities from the amide proton to H_ε(Lys) were readily observed after 8 scans were accumulated. Squared sine bell functions were used in both dimensions. The width of the functions was ca. 0.7 times t2 or t1, respectively, shifted by approximately $\pi/2$ in both dimensions.

The NOESY[25,26] experiments were carried out with the 90°-t1-90°-mix-90°-FID pulse sequence and 1.5 s solvent presaturation. A random variation of 10% was applied to the mixing time. In the case of the larger peptide (type II), NOESY at six different mixing times, namely 50, 150, 250, 350, 450 and 550 ms

were recorded without random delay, and 32 scans per t1-increment were acquired, resulting in negative nuclear Overhauser effects. For the smaller peptide (type III), the NOESY experiments were carried out with 40 scans per t1-increment at a mixing time of 200 ms. After data processing with $\pi/2$ shifted squared sine bell functions large numbers of correlations were observed.

The cross-peak volumes of the NOESY data sets obtained from the type II telopeptide were quantified at mixing times of 50, 150, 250, 350, 450 and 550 ms to check for linear build-up of the nuclear Overhauser effects. The cross-peak volumes were quantified in the following way: (i) cross-peaks were integrated and carefully checked for contributions from baseline distortions, such as t1 and t2 ridges, by integrating and, if necessary, subtracting the integral of nominally empty areas perpendicular to the actual cross-peaks; (ii) whenever possible, the volume was measured in both cross-peaks, above and below the diagonal, and the two values were averaged to minimize slight phase imperfections and other minor artifacts; (iii) where applicable, the volumes were normalized for the contributions of more than two protons to the cross-peak by a factor $(2N_aN_b)/(N_a+N_b)$, where N_a and N_b are the number of protons responsible for the diagonal peaks a and b, respectively.[27] (iv) for the type II telopeptide the $\beta\beta'$ interactions of residues Asp^5, Asp^{20} and Tyr^{24} were used as reference (1.8 Å); for the type III telopeptide the $\beta\beta'$ interactions of the Tyr 17, 19 and 21 moieties were used as a reference (1.8 Å) to calculate the inter-proton distances. It was not possible to establish the degree of attenuation of the exchangeable amide protons in a precise manner because of spectral overlap. Qualitatively speaking however, the loss of intensity was small, as evidenced by the few cases where a precise integration and subsequent comparison with non-exchangeable reference protons was possible. The absence of correction

factors for the exchangeable protons renders the distance constraints longer, so that no unnaturally close proximities are created.

Results

Circular dichroism spectroscopy

The comparison of CD spectra of the type II collagen C-telopeptide in two solvents is presented in Figure 3.1 from which it can be seen that the peptide has a virtually disordered conformation in water as solvent. However in CH_3OH/H_2O (80/20), it adopts a degree of ordered conformation, the CD profile of which is given in terms of the difference spectrum displayed in the inset to the figure. The features of the difference spectrum suggest, perhaps, 8% α-helix based on the relationship

$$\% \ \alpha\text{-helix} = - \ (\text{Mean Residue Ellipticity at 222 nm}) / 29700,$$

where - 29700 deg cm^2 $dmol^{-1}$ is the estimated value of the mean residue ellipticity for 100% helix formation of a sequence of 10 amino acid residues using the rationale and equation 5 of Chang et al.[28] A subsidiary conclusion is that the peptide is not aggregated in CH_3OH/H_2O (80/20), since decreasing the peptide concentration 5-fold had no significant effect on the CD spectrum.

NMR spectroscopy

The resonance assignments (Tables 3.1 and 3.2) of both telopeptides are based on TOCSY experiments[21] and sequential NOEs of the type α(i)-NH(i+1), NH(i)-NH(i+1) and β(i)-NH(i+1).[29]

Type II collagen α-1 chain C-telopeptide

The assigned resonances of the type II collagen α-1 chain C-telopeptide are summaried in Table 3.1, which also includes the temperature dependence

coefficients and coupling constants of the NHs (where available), and the relevant section of the TOCSY recorded with a mixing time of 80 ms shown in Figure 3.3. Vicinal coupling constants of the type $^3J_{\alpha NH}$ were measured in the one-dimensional spectra, at different temperatures as required to minimize spectral overlap. Whenever possible, the values of the coupling constants were checked over the temperature range -5 to 25°C, and changes of more than 0.3 Hz were not found, an indication that the conformation does not change significantly. In all cases, the chemical shifts of the amide protons varied linearly with temperature with correlation coefficients better than 0.98, indicating that major conformational changes are not present in the temperature range under study. Several amide protons were observed to have smaller temperature dependence coefficients (below 3.0 ppb/K), which is a indication for a possible intramolecular hydrogen bond, or inaccessibility to solvent for steric reasons.

NOE build-up curves for several pairs of proton in the type II collagen α-1 C-telopeptide are depicted in Figure 3.3. One can see that, for the proton pairs whose distance could not change with the local conformation, the NOE intensities reached their highest point near the mixing time of 200 ms; while for ones whose distance could be changed with local conformation, the maximums were reached later near 450 ms. These differences could be caused by the combination of the rapid random motion of the molecule and the different correlation time at different parts of the molecule.

NOE information for the type II collagen α-1 chain C-telopeptide was extracted from a NOESY spectrum with 10% random delay and a mixing time of 150 ms. The observed effects are summarized schematically in Figure 3.5 (top). Basically, all the sequential NOEs, including NH(i)-NH(i+1), were observed along the chain, except where spectral overlap precluded measurements. Medium range and long range ($|i-j| >= 2$) NOEs, which are the most important indication

for the folded structure, were not found under careful examination, especially near the place where the lower temperature dependence coefficients were observed.

Type III collagen α-1 chain C-telopeptide

The assigned resonances of the type III collagen α-1 chain C-telopeptide are summarized in Table 3.2 and the relevant sections of the contour plot of the NOESY spectrum recorded at 200 ms mixing time are depicted in Figure 3.4. One can see that some amide-amide interactions overlap completely at 10°C. Recording the spectra at different temperatures does not solve this problem as the amide temperature coefficients are not significantly different. This is not surprising in view of the size of the molecule and its relatively unstructured nature. NOESY spectra were carefully examined for intraresidue (i = j), sequential (|i-j| = 1), medium range (|i-j| <= 4) and long range (|i-j| > 4) NOEs and the observed effects are summarized schematically in Figure 3.5 (bottom). Essentially an uninterrupted sequence of NH(i)-NH(i+1) cross-peaks was observed. The continuous NH/NH NOE pattern is seemingly interrupted in places where spectral overlap prevented a quantification of the cross-peaks, and of course, around the Pro^{14} moiety. From an overview of the spectroscopic data presented in Table 3.2 and Figure 3.5 it is evident that effects associated with a non-extended conformation are only found around residues 8 to 11.

The temperature coefficients of the amide protons measured over a 20°C range are included in Table 3.2. Linearity was observed in all cases which is a good indication that no major conformational changes took place. This conclusion is confirmed by the NH-H_α coupling constants (where available) which also did not change by more than 0.1 Hz over the entire temperature range. The temperature coefficient of the amide proton of Glu^{11} (2.0 ppb/°C) is

clearly indicative for shielding from solvent whereas all the other coefficients in the range from 3.9 to 6.0 ppb/°C demonstrate essentially unrestricted accessibility to solvent.

Discussion

The NMR data presented in Tables 3.1 and 3.2 and the graphic summary thereof (Figure 3.5) indicate that the conformation in solution of the two molecules under study is somewhat different. As indicated by the parameters used for secondary structure determination by NMR (nuclear Overhauser enhancements, amide proton exchange rates[30] and/or amide temperature coefficients), both peptides undergo conformational averaging on the NMR time scale. With a few exceptions,[31] this phenomenon has so far prevented *quantitative* conclusions about conformations of short linear peptides in solution. Unless the degree of conformational freedom is significantly reduced by disulfide bonds or binding to metal atoms,[32,33] the spectroscopic indicators of secondary structure[34,35] are population-weighted averages over all conformations present. The non-linearity of the averaging process[36] renders it essentially impossible to assess to what extent individual conformers are present. Only very extensive molecular dynamics calculations on time-scales of several nanoseconds can possibly solve this impasse.[37]

Type II collagen α-1 chain C-telopeptide

Based on the NMR experiments, the conformation of the type II C-telopeptide is mostly extended. In two regions where smaller temperature dependence coefficients for the NH protons were observed, which are one of the necessary indications of folded structures in small peptides, folded structures cannot be deduced because of the lack of other evidence. In the region of

residues Ser^7 to Ala^{10}, small temperature dependence coefficient for the NH proton of Ala^{10} and smaller coupling constant for the NH proton of Ala^8 are consistent with a β-turn structure, but the medium range $H_{\alpha}(i)$ - NH(i+1) NOE, generally regarded as the most reliable indication for turn formation[31], was not observed. In the C-terminal, smaller temperature dependence coefficients were found for three NH protons, but severe spectral overlaps prevented further conformational analysis over that region. These smaller temperature dependence coefficients can only be interpreted in terms of local shielding from solvent for sterical reasons (large hydrophobic side-chains).

The apparent 8% α-helix in CH_3OH/H_2O (80/20) may be interpreted in terms of an average taken over a set of interconverting subpopulations, some of which have a random structure while others have a degree of ordered conformation, although not necessarily helical. Alternatively, it is at least conceivable that on the average 92% of the molecules are completely random, while 8% are completely helical. This result is a confirmation of the result from the NMR study. Combining these two studies, we conclude that there might be a very small percentage of folded structure for the molecule, most likely in the C-terminal end and between residues 7 and 10.

Type III collagen α-1 chain C-telopeptide

The conformation of the type III C-telopeptide is mostly extended (Table 3.2 and Figure 3.5), except for the Gly^8 - Glu^{11} segment where some indications of turn formation were observed. First and foremost, a very low (in comparison to other values, see Table 3.2) temperature coefficient was observed for NH(Glu^{11}), indicating significantly reduced exchange with the bulk solvent. In view of the absence of any large hydrophobic side-chains around Glu^{11}, this is most likely due to hydrogen bonding. This conclusion is confirmed by a medium

range nuclear Overhauser effect between the α-proton of Ala^9 and NH(Glu^{11}). Unfortunately, the conformational analysis is severely hampered by four difficulties: (i) the presence of two Gly moieties in the putative turn limits the number of possible interactions drastically because of the absence of side-chain protons; (ii) vicinal NH-H_α coupling constants are basically meaningless for Gly as two α-protons are present; (iii) Gly^8 and Gly^{10} have identical H_α and NH chemical shifts making the quantification of relevant NOEs other than the medium range effect mentioned before impossible; and (iv) the chemical shift difference between NH(Gly^{10}) and NH(Glu^{11}) is so small, that any NOEs between the two protons would be too close to the diagonal to be measured. However, the observation of the medium range H_α(i) - NH(i+2) NOE, generally regarded as the most reliable indication for turn formation, makes the conclusion about the existence of a turn fairly certain. Although hydrogen bonding is not a prerequisite for β-turn formation[38], it appears to be the major stabilization factor in the present peptide. The decision whether the turn is type I or type II is difficult in view of the problems with spectral overlap. As the main difference between the two turns is the 180° rotation of the amide linkage between the second and the third amino acids of the turn (Ala^9 and Gly^{10}), few spectroscopic features are different in the two turns. The main differences are the very short distance H_α(Ala^9) - NH(Gly^{10}) of 2.2 Å in β–II (3.4 Å in β–I) and the extremely long distance NH(Ala^9) - NH(Gly^{10}) of 4.5 Å in β-II (2.6 Å in β-I, idealized turn values from reference 29). As both kinds of interactions are overlapped by other NOEs, no firm conclusion can be made based on this information alone. In all our studies to this date involving β-turns, we always found type I β-turns which is not surprising in view of their 2-3 fold predominance over type II.[39] In all those cases, it was found that the temperature coefficient of NH(i+2) was also reduced, although not as much as NH(i+3). This can be understood in view of the fact

that in the β-I turn the NH proton points towards the inside, whereas in the β-II case it is pointing to the outside of the turn where the exchange with the solvent would be unhindered. In the present case there is no reduction of the temperature coefficient of NH(Gly^{10}), and hence there is the potential of the turn being type II. When analyzing the sequence in terms of turn probabilities,[39] a high probability for a β-II turn at residues GAGE is found.

References

1. Leibovich, S. J. & Weiss, J. B. (1970), *Biochim. Biophys. Acta.* **214**, 445-454.
2. Helseth, D. L. Jr. & Veis, A. (1981), *J. Biol. Chem.* **256**, 7118-7128.
3. Capaldi, M. J. & Chapman,, J. A. (1982), *Biopolymers* **21**, 2291-2313.
4. Otter, A., Scott, P. G. & Kotovych, G. (1988), *Biochemistry* **27**, 3560-3567.
5. Otter, A., Scott, P. G. & Kotovych, G. (1987), *J. Am. Chem. Soc.* **109**, 6995-7001.
6. Otter, A., Kotovych, G. & Scott, P. G. (1989), *Biochemistry* **28**, 8003-8010.
7. Liu, X., Scott, P. G., Otter, A. & Kotovych, G. (1990), *J. Biomol. Struct. Dyn.* **8**, 063-080.
8. Otter, A., Scott, P. G. & Kotovych, G. (1993), *Biopolymers*, in press.
9. Su, M.-W., Lee, B., Ramirez, F., Machado, M. & Horton, W. (1989), *Nucleic Acids Res.* **17**, 9473.
10. Miller, E. J. (1984), *Extracellular Matrix Biochemistry*, Elsevier, New York, 41-81.
11. Lapiere, Ch. M., Nusgens, B. & Pierard, G. E. (1977), *Connective Tissue Research* **5**, 21-29.
12. Keene, D. R., Sakai, L. Y., Bächinger, H. P. & Burgeson, R. E. (1987), *J. Cell. Biol.* **105**, 2393-2402.
13. Fietzek, P. P., Allmann, H., Rauterberg, J., Henkel, W., Wachter, E. & Kuhn, K. (1979), *Hoppe-Seyler's Z. Physiol. Chem.* **360**, 809-820.
14. Chu, M.-L., Weil, D. DeWet, W., Bernard, M, Sippola, M. & Ramirez, F. (1985), *J. Biol. Chem.* **260**, 4357-4363.
15. Yamada, Y., Kuhn, K. & deCrombrugghe, B. (1983), *Nucleic Acids Res.* **11**, 2733-2744.
16. Jones, E. Y. & Miller, A. (1987), *Biopolymers* **26**, 463-480.

17. *Handbook of Biochemistry and Molecular Biology*, 3rd edition, Proteins, Volume I, p. 186 (Fasman, G.D., Ed.), CRC Press, Cleveland, Ohio.

18. Chen, G. C. & Yang, J. T. (1977), *Anal. Letters* **10**, 1195-1207.

19. Ernst, R. R., Bodenhausen, G. & Wokaun, A. (1987), *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, Clarendon Press, Oxford.

20. States, D. J., Haberkorn, R. A. & Ruben, D. J. (1982), *J. Magn. Reson.* **48**, 286-292.

21. Braunschweiler, L. & Ernst, R. R. (1983), *J. Magn. Reson.* **53**, 521-528.

22. Varian, NMR Spectrometer Systems, System Operation Manual, Publ. No. 87 -195100-00, Rev. A0689.

23. Otting, G., Widmer, H., Wagner, G. & Wüthrich, K. (1986), *J. Magn. Reson.* **66**, 187-193.

24. Bax, A. & Davis, D. G. (1985), *J. Magn. Reson.* **65**, 355-360.

25. Jeener, J., Meier, R., Bachmann, P. & Ernst, R. R. (1979), *J. Chem. Phys.* **71**, 4546-4553.

26. Kumar, A., Ernst, R. R. & Wüthrich, K. (1980), *Biochem, Biophys. Res. Commun.* **95**, 1-6.

27. Williamson, M.P. & Neuhaus, D. (1987), *J. Magn. Reson.* **72**, 369-375.

28. Chang, C. T., Wu, C.-S. C. Wu & Yang, J. T. (1978), *Anal. Biochem.* **91**, 13-31.

29. Wüthrich, K. (1986), *NMR in Proteins and Nucleic Acids*, Wiley Interscience, New York.

30. Wagner, G. (1983), *Quart. Rev. Biophys.* **16**, 1-57.

31. Dyson, H. J. & Wright, P. E. (1991), *Annu. Rev. Biophys. Biophys. Chem.* **20**, 519-538.

32. Lee, M. S., Gippert, G. P., Soman, K. Y., Case D. A. & Wright, P. E. (1989), *Science* **245**, 635-637.

33. Marsden, B. J., Hodges, R. S. & Sykes, B. D. (1989), *Biochemistry* **28**, 8839-8847.

34. Dyson, H. J., Rance, M., Houghten, R. A., Lerner, R. A. & Wright, P. E. (1988), *J. Mol. Biol.* **201**, 161-201.

35. Dyson, H. J., Rance, M., Houghten, R. A., Wright, P. E. & Lerner, R. A. (1988), *J. Mol. Biol.* **201**, 201-217.

36. Jardetzky, O. & Roberts, G. C. K. (1981), *NMR in Molecular Biology*, Academic Press, New York.

37. Brüschweiler, R., Blackledge, M. & Ernst, R. R. (1991), *J. Biomol. NMR* **1**, 3-11.

38. Rose, G. D., Gierasch, L. M. & Smith, J. A. (1985), *Adv. Protein. Chem.* **37**, 1-109.

39. Wilmot, C. M. & Thornton, J. M. (1988), *J. Mol. Biol.* **203**, 221-232.

Table 3.1 ^1H Chemical Shifts[a], Coupling Constants[b] and NH Temperature Coefficients[c] of the Type II Collagen α-1 Chain C-Telopeptide[d]

	NH	H_α[e]	H_β[e]	H_γ[e]	other H	$J_{NH-\alpha}$	NH temp. coeff.
G^1	8.38	4.17				5.4/5.4	7.0
		3.97					
P^2	-	4.38	2.23,2.23	2.07,1.94	δ: 3.63 3.58		---
G^3	8.73	3.96				6.0/6.0	5.6
		3.81					
I^4	7.94	4.18	1.84	1.43,1.11	δ:0.82		4.2
				0.89(CH_3)			
D^5	8.53	4.61	2.82,2.64			7.0	5.8
M^6	8.68	4.62	2.25,2.00	2.67,2.53	ε: 2.00	6.4	6.2
S^7	8.65	4.18	3.96,3.91			5.0	3.4
A^8	8.33	4.09	1.18			4.6	6.6
F^9	7.92	4.42	3.27,2.97		7.25		4.2
A^{10}	7.89	4.23	1.44				2.2
G^{11}	8.18	3.93				6.0/6.0	4.7
		3.87					
L^{12}	8.00	4.43	1.74,1.74	1.66	0.93,0.87	8.0	3.5
G^{13}	8.31	4.07				6.0/6.0	5.6
		4.02					
P^{14}	-	4.38	2.23,2.23	2.07,1.94	δ: 3.64 3.58	---	---
R^{15}	8.52	4.28	1.88,1.76	1.64,1.64	δ: 3.18 3.18	7.0	6.3
					ε:NH: 7.53		
E^{16}	8.41	4.26	2.04,2.04	2.32,2.32		6.3	2.3
K^{17}	8.25	4.34	1.87,1.70	1.46,1.46	δ: 1.62 1.62	7.0	4.6
					ε: 2.90 2.90		
G^{18}	8.21	4.11				6.5	4.8
		4.04					
P^{19}	-	4.41	2.23,2.23	2.07,1.99	δ: 3.63 3.56	---	-
D^{20}	8.56	4.89	2.86,2.58			6.4	7.9
P^{21}	-	4.27	2.36,2.36	2.03,1.92	δ: 4.00 3.96	---	-

L^{22}	8.39	4.26	1.82,1.64	1.55	δ:0.95,0.87	7.0	1.9
Q^{23}	7.80	4.02	2.18,2.18	2.03,2.03	δNH_2: 7.73,6.93	5.4	0.8
Y^{24}	7.82	4.45	3.12,2.93		2,6: 7.10 3,5: 6.72	6.7	4.6
M^{25}	7.90	4.43	2.46,2.38	2.10,1.97	ε:2.00		2.7
R^{26}	7.90	4.37	1.88,1.76	1.63	δ: 3.16 3.16 NH: 7.60		3.3
A^{27}	7.94	4.16	1.34				5.3

a) Chemical shifts are reported in ppm relative to the methyl group of CD_3OH at 3.30 ppm downfield from TMS.

b) Coupling constants in Hertz. Whenever possible, they were measured in one-dimensional spectra. When it was not available from one-dimensional spectra, it remained unknown.

c) The temperature coefficients of the amide protons are listed in - ppb/°C. They are the result of a linear regression analysis of the chemical shifts measured at 0.0, 5.0, 10.0, 15.0, 20.0, and 25.0°C and accurate to within +/- 0.3 ppb/°C.

d) The data were recorded at 500 MHz in CD_3OH/H_2O (80/20 v%) at a concentration of 2.1 mM. The pH was 5.4 and the temperature was 0.0 ± 0.1°C.

e) Pairs of geminal protons x (x = β, γ, δ, ε and α(G)) are not assigned stereospecifically. When two distinct lines were observed, the larger chemical shift was arbitrarily assigned to x and the smaller to x'.

Table 3.2 ^1H Chemical Shifts[a], Coupling Constants[b] and NH Temperature Coefficients[c] of the Type III Collagen α-1 Chain C-Telopeptide[d]

	NH	H_α [e]	H_β [e,f]	H_γ [e]	other H	$J_{NH-\alpha}$	NH temp. coeff.
G^1							
G^2							
G^3	8.42	3.97					
		3.92					
V^4	8.15	4.03	2.11	0.96,0.96		6.5	5.7
A^5	8.41	4.29	1.39			6.0	5.8
S^6	7.98	4.37	3.88,3.81			7.0	3.9
L^7	8.17	4.33	1.71,1.69	1.60	δ: 0.91 0.86	7.0	3.9
G^8	8.42	3.92					
		3.92					
A^9	8.27	4.31	1.39			5.5	4.6
G^{10}	8.42	3.92				2x5.8	5.4
		3.92					
E^{11}	8.46	4.25	2.07,1.99	2.28,2.28		6.0	2.0
K^{12}	8.33	4.37	1.89,1.72	1.44,1.44	δ: 1.63 1.63	7.5	5.0
					ε: 2.92 2.92		
G^{13}	8.12	4.02				2x5.4	4.1
		4.02					
P^{14}	---	4.40	2.13,1.94	1.94,1.94	δ: 3.55	---	---
V^{15}	8.21	4.08	2.03	0.91,0.88		7.2	6.0
G^{16}	8.40	3.92				2x5.5	5.5
		3.74					
Y^{17}	8.27	4.39	3.02,2.85		2,6: 7.03	6.0	5.4
					3,5: 6.71		
G^{18}	8.56	3.96				2x6.0	4.7
		3.61					
Y^{19}	8.02	4.43	3.02 2.85		2,6: 7.01	7.5	5.6
					3,5: 6.70		
E^{20}	8.77	4.19	1.91,1.83	2.16,2.16		6.0	5.3

Y^{21}	8.02	4.53	3.08	2.85	2,6: 7.05 3,5: 6.73	7.5	4.3
R^{22}	7.69	4.18	1.85,1.68	1.54,1.54	δ: 3.14 3.14 ε: 7.48	7.7	5.0

a) Chemical shifts are reported in ppm relative to the methyl group of CD_3OH at 3.30 ppm downfield from TMS.

b) Coupling constants in Hertz, measured in one-dimensional spectra.

c) The temperature coefficients of the amide protons are listed in ppb/°C. They are the result of a linear regression analysis of the chemical shifts measured at 0.0, 10.0 and 20.0°C.

d) The data were recorded at 500 MHz in CD_3OH/H_2O (80/20 v%) at a concentration of 3.0 mM. The pH was 5.5 and the temperature 10.0 ± 0.1 °C.

e) Pairs of geminal protons x (x = β, γ, δ, ε and α(G)) are not assigned stereospecifically. When two distinct lines were observed, the larger chemical shift was arbitrarily assigned to x and the smaller to x'.

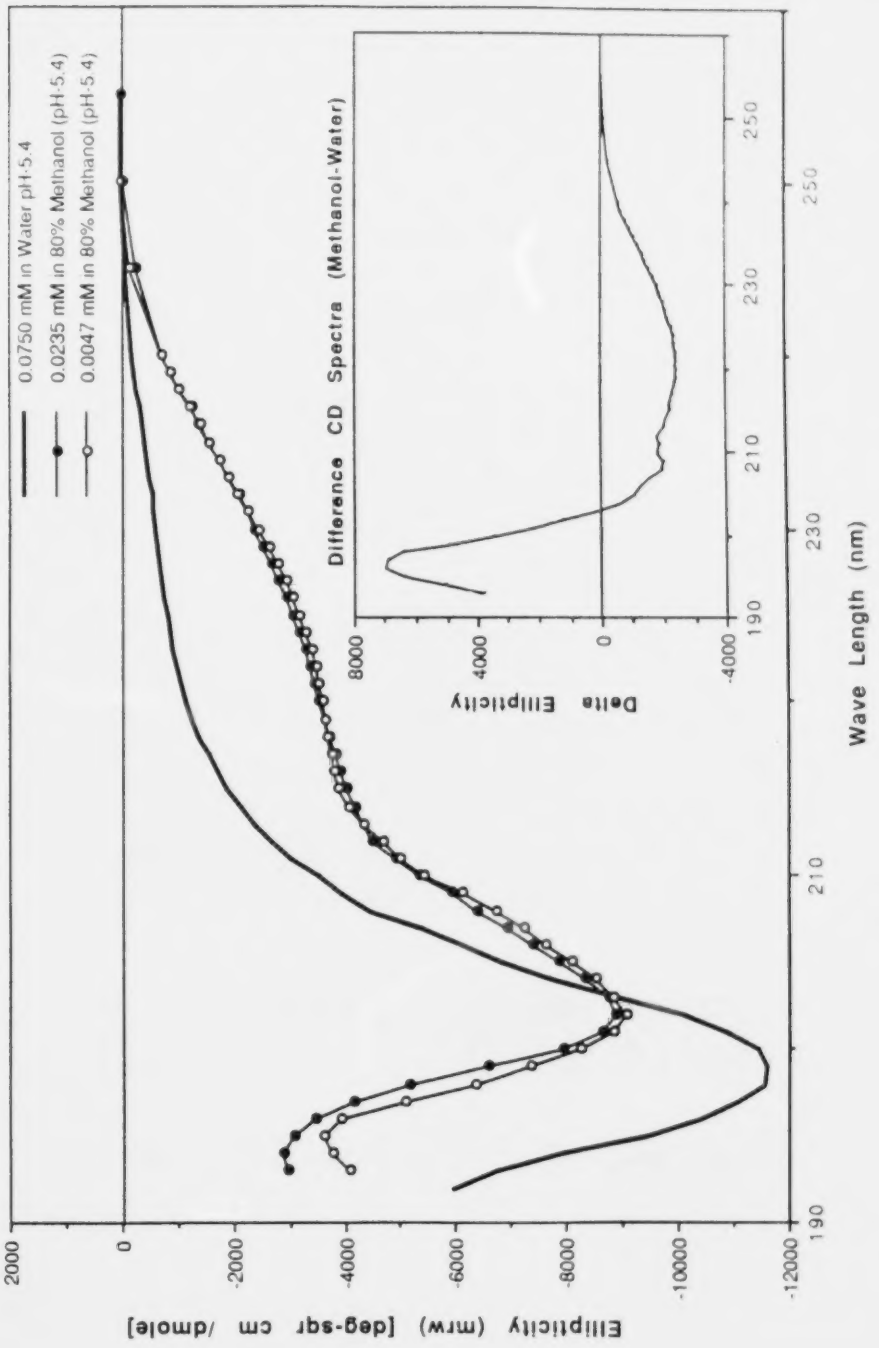
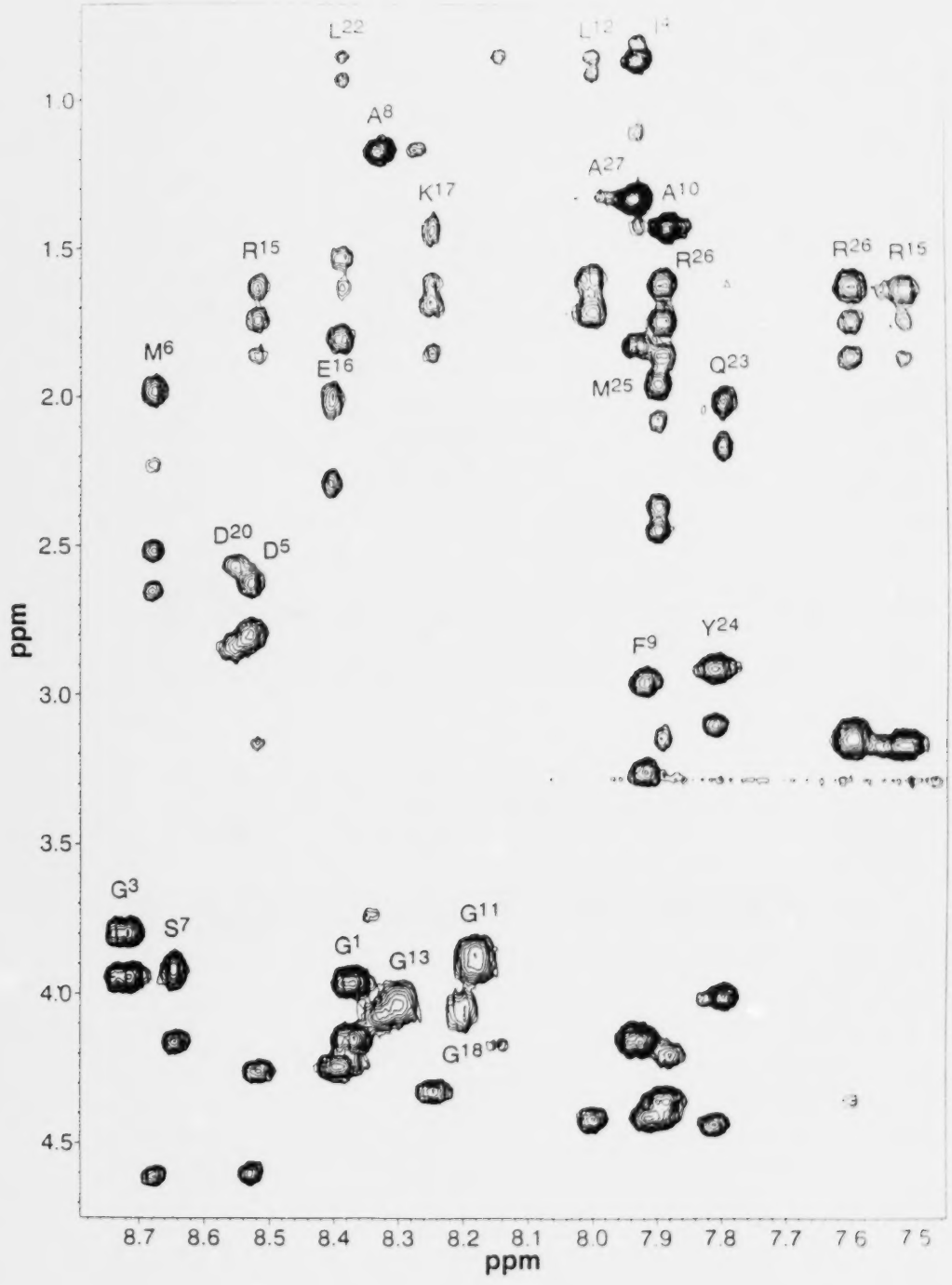


Figure 3.1 Circular Dichroism spectra of the type II collagen α-1 chain C-telopeptide at 10 °C. For experimental details see "Materials and Methods".

Figure 3.2 The contour plot of a phase-sensitive 500 MHz TOCSY experiment of the type II collagen α-1 chain C-telopeptide at a mixing time of 80 ms. The concentration of the sample was 2.1 mM at pH 5.4 in CD_3OH/H_2O (80:20), temperature 0.0°C.


L22
L12
I4
A8
A27
A10
K17
R15
R26
R26
R15
M6
E16
M25
Q23
D20
D5
F9
Y24
G3
S7
G1
G13
G11
G18
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
ppm
8.7
8.6
8.5
8.4
8.3
8.2
8.1
8.0
7.9
7.8
7.7
7.6
7.5
ppm

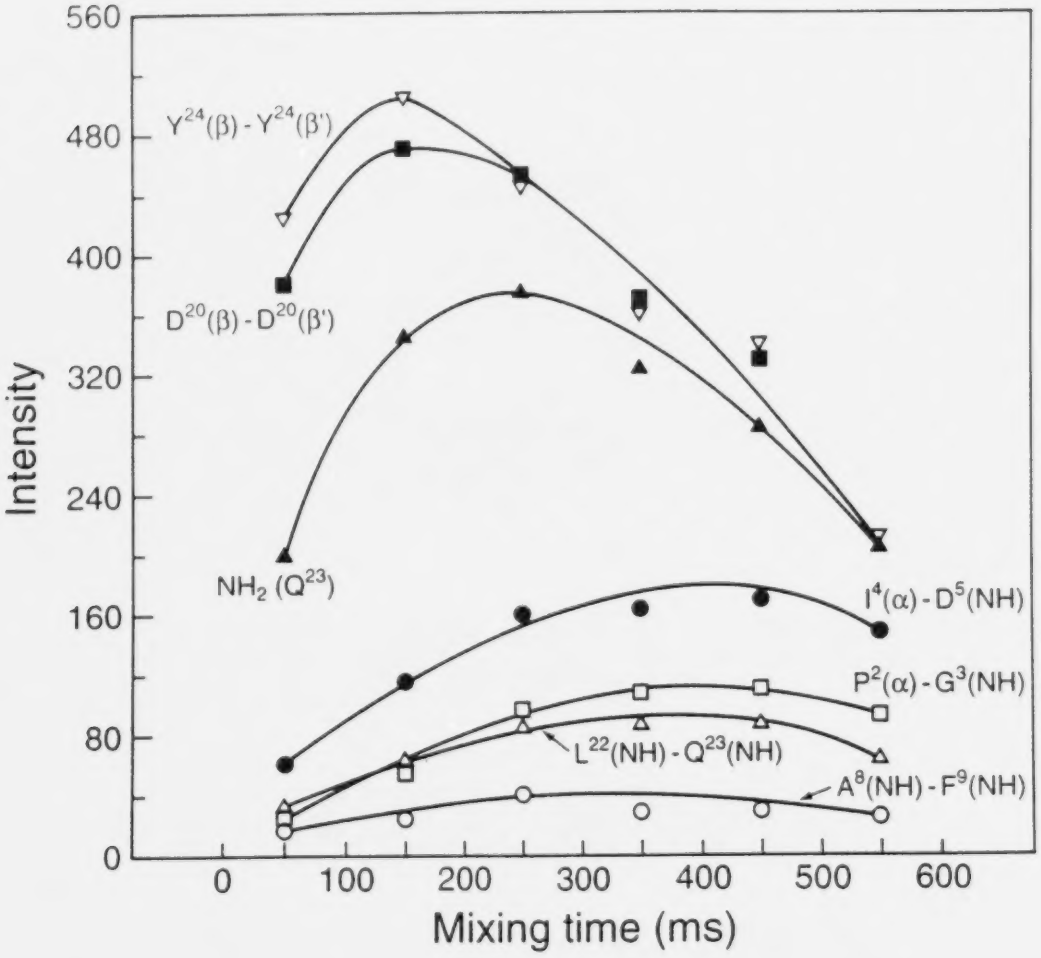


Figure 3.3 Plot of several build-up curves of the pairs of protons in the type II α-1 C-telopeptide indicated in the plot. The concentration of the sample was 2.1 mM at pH 5.4 in CD_3OH/H_2O (80:20), temperature 0.0°C. NOESY without the random delay at different mixing times of 50, 150, 250, 300, 350, 400, 450 and 550 ms were used. These experiment data were processed with identical parameters. The number for the intensity is arbitrary. See text for detail.

Figure 3.4 Two sections of a 500 MHz phase-sensitive NOESY experiment of the type III collagen α-1 chain C-telopeptide at a mixing time of 200 ms. The concentration was 3.0 mM at pH 5.5 in CD_3OH/H_2O (80:20), temperature 10.0°C. Bottom: the NH/NH region; top: H_α/NH region (for clarity, only the peaks pertinent to the β-turn conformation around residues G^8 to E^{11} are labeled).

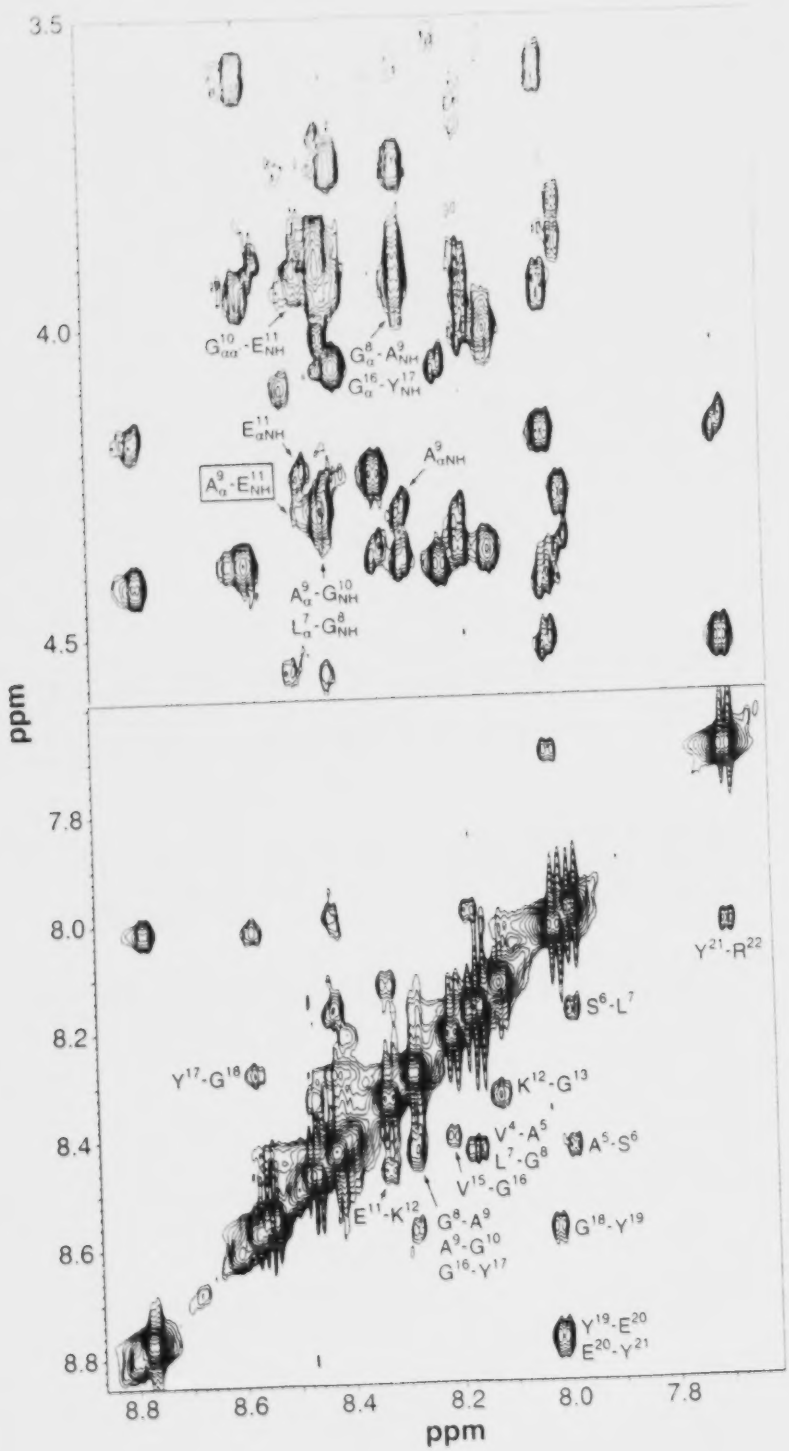
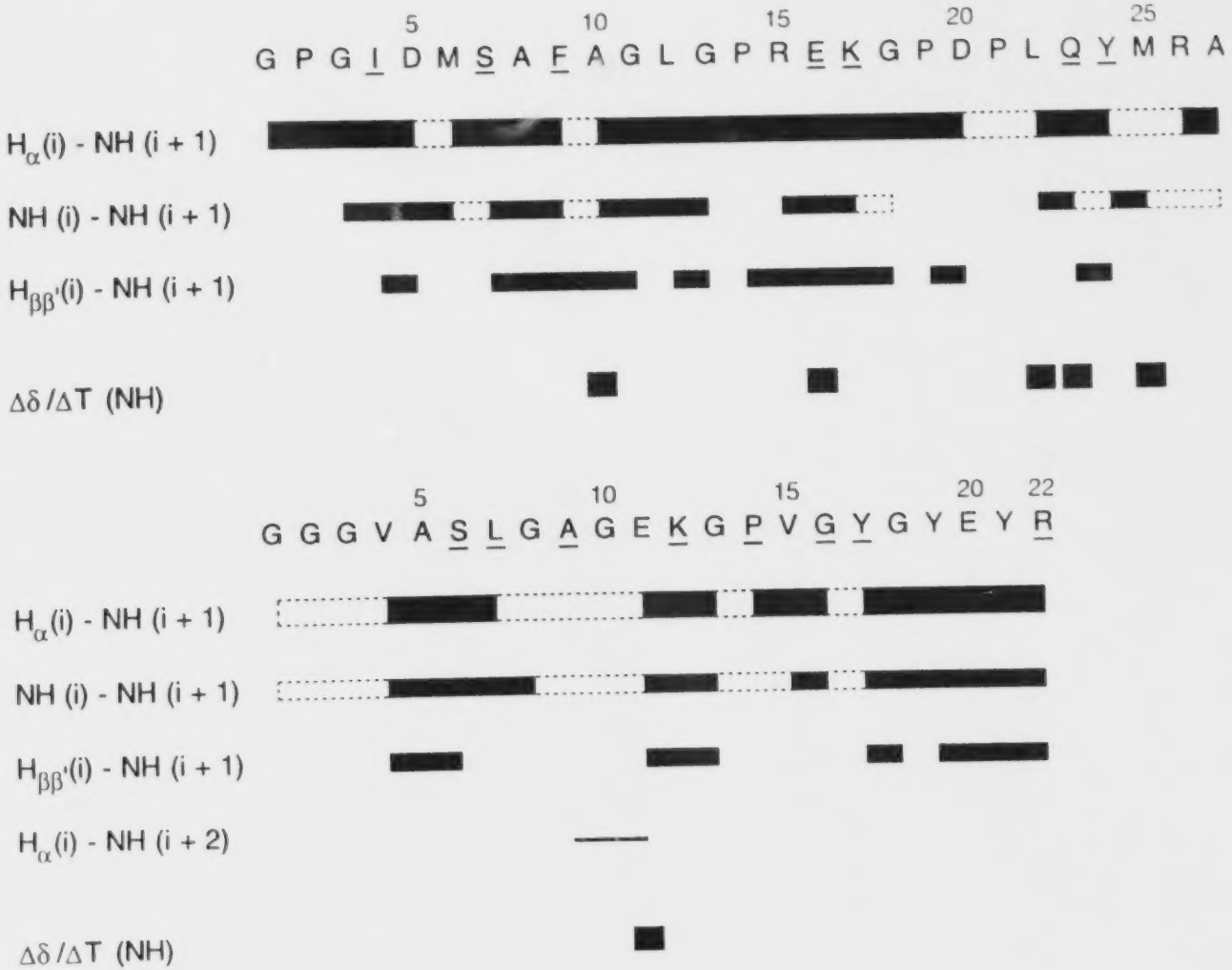

3.5
4.0
4.5
ppm
$G^{10}_{\alpha\alpha}$-E^{11}_{NH}
G^{8}_{α}-A^{9}_{NH}
G^{16}_{α}-Y^{17}_{NH}
$E^{11}_{\alpha NH}$
$A^{9}_{\alpha NH}$
A^{9}_{α}-E^{11}_{NH}
A^{9}_{α}-G^{10}_{NH}
L^{7}_{α}-G^{8}_{NH}
7.8
8.0
8.2
8.4
8.6
8.8
ppm
Y^{21}-R^{22}
S^{6}-L^{7}
Y^{17}-G^{18}
K^{12}-G^{13}
V^{4}-A^{5}
L^{7}-G^{8}
A^{5}-S^{6}
V^{15}-G^{16}
E^{11}-K^{12}
G^{8}-A^{9}
A^{9}-G^{10}
G^{16}-Y^{17}
G^{18}-Y^{19}
Y^{19}-E^{20}
E^{20}-Y^{21}

Figure 3.5 The amino acid sequences of the type II (top) and type III (bottom) collagen α-1 chain C-telopeptides with unique residues underlined. The observed sequential and medium range NOEs are indicated schematically by solid bars. The height of the bars is an approximation for the strength of the NOEs. Where proline moieties are present, the $H_{\alpha}(i)$ - $H_{\delta\delta'}(i+1)$ instead of the $H_{\alpha}(i)$ - NH(i+1) interactions are listed. NOEs shown with dashed lines are likely to be present but could not be quantified due to almost identical chemical shifts of the protons. The temperature coefficients of the amide protons are shown with a solid square for values below 3.0 ppb/°C.


5
10
15
20
25
G P G I D M S A F A G L G P R E K G P D P L Q Y M R A
H_α(i) - NH (i + 1)
NH (i) - NH (i + 1)
$H_{\beta\beta'}$(i) - NH (i + 1)
Δδ/ΔT (NH)
5
10
15
20
22
G G G V A S L G A G E K G P V G Y G Y E Y R
H_α(i) - NH (i + 1)
NH (i) - NH (i + 1)
$H_{\beta\beta'}$(i) - NH (i + 1)
H_α(i) - NH (i + 2)
Δδ/ΔT (NH)

CHAPTER 4

A Sequence-Dependent ^{1}H NMR Study on the Formation of β-Turns in Tetrapeptides Containing Charged Residues*

Introduction

The molecular mechanism involved in the folding of peptides and proteins remains one of the important questions in structural biology[1]. Recent studies[2-5] clearly demonstrate the ongoing interest in the factors favoring formation and stabilization of secondary structures at an early stage of the folding process. In particular, the observation of a stable β-turn in short linear peptides is important since these structures have been postulated to play a significant role in initiating protein folding[6-8]. A turn is an important structural feature in peptides and proteins and is the site in the polypeptide chain where the polypeptide direction is reversed[9,10]. In the case of the β-turn, a hydrogen bond may be formed between the carbonyl oxygen of residue i and the amide proton of residue i+3. Even in short linear peptides, a beta turn can be stable, as observed for the pentapeptide Tyr-Pro-Gly-Asp-Val which is 50% beta turn at 5°C[2]. This conformation is in equilibrium with an extended conformation in water. The authors propose that it is the local amino acid sequence in the folded protein that determines the β-turn conformation because it is observed in short pentapeptides. Consequently, such a sequence is sufficiently stable so that it can initiate the direction of the folding pathway. The conformational space is significantly restricted by reverse turns in the folding polypeptide chain and this

---

* A version of this chapter has been published. Liu, X., Scott, P. G., Otter, A. & Kotovych, G. (1992), *Biopolymers* **32**, 119-130.

allows distant sections of a polypeptide to come into close proximity[2,3].

In the present study, the importance of side-chain charge interactions in the formation of β-turns is studied. The tetrapeptide NAc-DEKS-NH_2 has been previously studied[11]. It forms a type I beta-turn which has been observed in the isolated tetrapeptide as well as in the type I collagen α-1 chain N-telopeptide[12]. To study the importance of the side chains (their length, ionization state and the order within the sequence), the following protected NAc-tetrapeptide amides were synthesized: NEKS, EEKS, DDKS, DQKS, NQKS, DERS, NERS, EERS, DDRS, NDRS, DQRS and DKES. In addition, based on the statistical predictions of Wilmot and Thornton[13], the peptides NPDM, NSDM and NDDS were studied.

Materials and Methods

The tetrapeptides in this study consist of two series. The first series is related to our previous study of the type I collagen α-1 chain N-telopeptide (16-mer)[12] and the tetrapeptide DEKS[11] and includes NEKS, EEKS, DDKS, DQKS, NQKS which have the same residues at positions 3 and 4 as in DEKS, but have different side-chains and charges for residues at positions 1 and 2; DERS, NERS, EERS, DDRS, NDRS, DQRS where K has been replaced by R at position 3; and DKES which is the variant of DEKS obtained by switching the second and third residues. The second series include NSDM, NPDM and NDDS. These molecules were included in the present studies in view of the results of a statistical analysis and predictions of β-turns in proteins[13]. The sequence NPDM is composed of the four amino acids that are among the most often found in positions 1 through 4 in type I β-turns, and therefore could potentially serve as a model compound for the present investigation. Because of the absence of an NH proton at the important position 2 in NPDM, we included NSDM, the next probable β-turn forming tetrapeptide in the sequence. NDDS was selected since

it has the highest probability of forming a type I β-turn without the possibility of attractive charge interactions. All peptides were synthesized as the N-acetyl-amides. DEKS was from Alberta Peptide Institute, University of Alberta. NQKS, NPDM and NSDM were from Multiple Peptide Systems, San Diego, CA. All other peptides were synthesized on an LKB Biolynx^{TM} 4175 solid-phase peptide synthesizer using Fmoc-amino acid derivatives (Pharmacia (Canada) Ltd, Baie d'Urfe, Quebec). All peptides were purified by reversed-phase HPLC on a Vydac TP218 column (The Separations Group, Hesperia, CA). Purity was confirmed by determination of amino acid composition.

The NMR samples were prepared by dissolving the purified solid material in ca. 0.2 mL of water (H_2O). The pH was then adjusted by adding small aliquots of diluted NaOH or HCl, followed by adding the rest of the solvent, which consisted of 170 μL CD_3OH and 130 μL CD_3OD to make the solvent a 60/40 methanol/water mixture. Before the NMR experiments, the samples were degassed with argon. The concentrations of the samples were between 3.9 and 6.8 mmol/L and the pH of the samples was in the range of 5.5 to 6.6. Whenever the same peptide was studied at different pH, the second sample was prepared in the same way. For several samples, spectra at different concentrations were compared and the chemical shifts were found to be independent of the concentration, indicating the absence of aggregation or other structural changes.

All experiments were carried out on Bruker AM300 and AM400 NMR spectrometers at ambient temperature (295K), except for the temperature studies. Three probes were used, namely a broadband probe on the AM300, a selective ^1H probe and an inverse probe on the AM400. Since the inverse probe has a much higher sensitivity than the other two, it was used to carry out all of the ROESY experiments as well as experiments on the more dilute samples. The temperature studies were all carried out on the AM400. The temperature

dependence of the amide protons was calculated by comparing spectra recorded at different temperatures in the range from 283K to 300K and a maximum variation of +/- 0.3 ppb/K was observed for the temperature coefficients. Each temperature was calibrated by inserting a thermocouple inside an NMR tube, filled with the same mixed solvent, into the magnet before and after the experiments. When the temperature dependence was calculated, the calibrated and averaged value for each temperature was used. For 5 peptides (DERS, NEKS, DDRS, DDKS, DKES) both the selective and the inverse probe were used for the temperature study, and the results are averaged values over 8 temperatures. For the remainder of the peptides the results were obtained from spectra at 4 temperatures using only the inverse probe.

The FID acquisition was controlled by an ASPECT 3000 computer using 1987 or 1989 Bruker DISNMR software. For data processing the same computer and software were used, except that the Fourier transformations were carried out on the ASPECT 1000 computer with an array processor. The chemical shifts were measured relative to CD_3OH (Merck, Sharp & Dome, MD-67, 99.2% D) at 3.30 ppm (^1H).

One-dimensional proton spectra were recorded with a sweep width of 10 ppm and 16K (zero-filled to 32K) data points. The presaturation of the solvent peak(s) was performed during a 3s relaxation delay at the lowest possible decoupler power level to avoid saturation transfer and possible wipe-out of the exchangeable protons. The frequency of the decoupler was changed every 0.3s between the two decoupler positions, one at the OH-signal and another at the signal of the undeuterated fraction of CD_3OH at 3.30 ppm. To account for the big difference in intensity between the two solvent resonances, the former signal was saturated twice as long as the methyl group. Usually a total of 64 scans were accumulated, and the resulting FID's were resolution enhanced by means

of a Gaussian multiplication with a line broadening factor of typically -3 Hz.

Two-dimensional NMR experiments, i.e. correlated spectroscopy (COSY)[14] and spin-locked NOE (ROESY or CAMELSPIN)[15,16], were employed for the assignment of the peptide spectra and for the detection of spectral features related to secondary structure. An example of the parameters used for the two-dimensional NMR techniques is given in Table 4.1. For different molecules, certain parameters such as sweep width, acquisition time and t_1-increment were slightly different, since those parameters depend on the field strength of the spectrometer and were adjusted slightly in order to achieve the maximum resolution in both dimensions. The two-dimensional experiments were all carried out in the phase-sensitive mode using time-proportional phase increments(TPPI)[17]. Solvent suppression was achieved as described for the one-dimensional spectra except that the relaxation delay was 1.8s instead of 3s. In addition, continuous irradiation was applied to the water signal at all times except during acquisition[18]. The phase constants were determined by using the second FID, which was multiplied by the weighting function, Fourier transformed, and then phased. The phase constants thus determined were applied to all spectra in the t_2 domain. No phase correction was done in t_1 except for the ROESY experiments where a basic $\pi/2$ phase correction was required. The weighting functions for both dimensions were sine-bell window functions. Shifted sine-bell functions were applied to obtain the best results in the "fingerprint" region of a COSY spectrum; on the other hand unshifted sine-bell functions were the better choice for the high-field regions of COSY spectra. The resulting two-dimensional matrix was usually slightly rephased in order to obtain the best possible phase in the area of interest. The two-dimensional proton spectra suffered somewhat from the appearance of t_1-noise and ridges[19,20], but these problems were not serious enough to affect our analysis.

In the ROESY experiments the same conditions that were used in our previous study of the tetrapeptide DEKS[11], were repeated here. The spin-lock conditions were achieved by a series of hard pulses of small flip angle[21] and the spin locking time was set to 300 ms at an average field strength of 4.0 kHz which was found to be suitable for peptides of similar size in previous studies[22]. Since ROE cross-peaks show a non-negligible offset dependence[23], they often cannot be observed at the edge of a spectrum, irrespective of the internuclear distances involved. This is especially true for NH to NH interactions which are extremely important for the determination of secondary structures in peptides and proteins[24]. Therefore, a ROESY experiment was performed with the transmitter set in the middle of the NH region if the NH-NH interactions were not observed with the transmitter set at ca. 5 ppm. When NH-NH interactions were observed, then the ROESY experiments were carried out at two additional mixing times (100 ms and 200 ms). Usually the ROESY spectra showed some baseline distortions around intense diagonal peaks. Therefore, the spectra were subjected to a baseline correction routine which fits a polynomial function of fifth degree to the baseline (ABS baseline correction), and the problem was eliminated almost completely in both dimensions. It is noteworthy that the ROESY experiments were almost free of any spurious resonances (identified by their phase relative to the diagonal) due to magnetization transfer between scalar coupled spins[25,26]. In the so-called "fingerprint" region[27] no such effects were present, indicating suitable spin-locking conditions[21].

Finally, the ROESY spectra were re-run on DEKS, DERS and NEKS at 500 MHz on a Varian UNITY 500 NMR spectrometer, on which better sensitivity is available. The samples for these experiments (5 mM, pH 5.8) were prepared in the same way as described above. Spin-locking was achieved by a series of 30° hard pulses, with two hard 90° pulses added at both ends of the spin-locking

period to compensate for the offset effect of the cross-peaks[23]. The strength for the spin-lock was set to 5.0 kHz and the spin-locking time to 300 ms. The sweep-width of the spectra was 5000 Hz and the transmitter was set on the water signal (ca. 5 ppm). Solvent suppression was achieved by presaturation during a 2s relaxation delay. The data were processed by using VNMR 3.1 software which was run either on the SUN 4/330 or SUN 4/470 computer. The ROESY experiments of the three peptides were run and processed with exactly the same parameters in order to estimate the ROE intensities qualitatively. The plot of the NH-NH region for each of the three molecules is shown in Figure 4.1.

Results

(a) ^{1}H resonance assignments

The spectral analysis of the tetrapeptides is based on the results of the phase-sensitive COSY[17] and ROESY[15,16] experiments. All spectra used for assignments were recorded with the transmitter set approximately in the middle of the spectral range (ca. 5 ppm). Since there was rarely spectral overlap in the two-dimensional spectra, the COSY experiments, combined with the sequential Hα(i)-NH(i+1) ROEs[24] in the ROESY experiments, allowed the complete assignment of all the protons of the tetrapeptides without any difficulty.

The assignments for all the peptides are summarized in Table 4.2, which also includes the results of a temperature study of the exchangeable protons for each tetrapeptide at nearly neutral pH together with NH-Hα coupling constants, which were obtained directly from the one-dimensional spectra. The data for six tetrapeptides that were studied at low pH are also presented in Table 4.2.

(b) Temperature coefficients of amide proton resonances

The temperature dependence of the chemical shifts of the amide protons was determined over the range from 284K to 300K. In all cases, the chemical shifts varied almost linearly with temperature, indicating that major conformational changes are not present in the temperature range under study. It is believed that the temperature dependence of an amide proton reflects differences in the extent to which the backbone amide proton participates in a hydrogen bond, though not necessarily in a β-turn conformation. It is noticeable that the values for the backbone amide proton resonance of residue 4, which was Ser in most molecules, vary greatly between peptides, from 2.7 in DERS to 6.4 in NDRS. A number of low temperature coefficients (below 4.0 ppb/K), which is indicative of shielding from the solvent, were observed, mainly for DERS, NEKS, DDRS, and DKES, probably because of hydrogen bonding. However, it has to be pointed out that since the temperature dependence of amide protons strongly depends on the solvent and the pH value, any quantitative comparison is difficult. The fact that small peptides in solution possess several conformations in equilibrium further aggravates this.

(c) pH dependence of amide protons in selected tetramers

In addition to DEKS, six tetrapeptides were investigated at pH values below the pKa of the Asp or Glu carboxyl groups. Upon lowering the pH, two interesting effects were observed for DERS and NEKS : (i) the chemical shifts of NH(E^2) shifted to high field by more than 0.5 ppm; (ii) The NH-Hα coupling constant of E^2 increased from 5.0 to 6.6 and 7.0, respectively. For other peptides we observed that the chemical shifts of NH(E) in the second and third position were subjected to a relatively large high-field shift of 0.3 ppm (EEKS and DKES) but only 0.1 ppm for the NH(E) in the first position (EEKS). It is

interesting that the chemical shifts for NH(E)s in the second position were on the low field side of 8.7 ppm (EERS, EEKS, DERS, NEKS, NERS), while those for NH(E)s in the first or third position were around 8.5 ppm (EERS, EEKS and DKES). In comparing the chemical shifts of NH(E) and NH(D), NH(D) in the first position shifted to low field upon lowering the pH (DERS, DKES, DDRS and DDKS), and NH(D) at the second position shifted in the same direction but less in extent (DDRS and DDKS).

(d) Use of the nuclear Overhauser effect to define conformations in solution

Conventional NOESY spectra are inappropriate since the peptides of interest are of such a size that cross-peaks are either nonexistent or too weak to be reliably observed because the correlation time of these molecules at 300 and 400 MHz is close to the rate at which the maximum possible NOE passes through zero[28]. An alternative approach for small molecules is the rotating frame NOE (ROESY) experiment. This experiment involves the use of a spin-lock following frequency labeling of a length sufficient to allow incoherent magnetization transfer between nuclei which are close together in space. The ROE remains positive irrespective of the value of the correlation time and the Larmor frequency. ROE cross-peaks appear as peaks with opposite intensity with respect to the diagonal peaks in the ROESY spectrum. The experiment must be used and interpreted with care, however, since artifacts can easily arise due to the presence of Hartmann-Hahn magnetization transfer[16,21,23,25,26]. Such artifacts can be detected by altering the conditions under which the experiment is carried out, for example, by changing the frequency offset of the carrier, the field strength or the mixing time of the spin-lock. In the present study, Hartmann-Hahn magnetization transfer was only observed for the two protons of the terminal NH_2 group or NH_2 groups in the side chain. But the

appearance of these cross-peaks does not affect the analysis in any way.

The cross-peaks between Hα(i) and NH(i+1), NH(i) and Hα(i), and NH(i) and Hβ(i), which show connectivities between consecutive amino-acid residues or within the same residue, were observed for all the peptides. Besides these, well-defined ROEs between NH(i) and NH(i+1) were observed only for DERS and NEKS, and the patterns were the same as were found for DEKS, except the intensities were different. Although not used in any quantitative sense, ROE intensities were estimated qualitatively from two-dimensional contour plots and relative to the generally strong sequential cross-peaks.

For DERS and NEKS, for which we observed NH-NH interactions in the ROESY spectra, we took particular care that the ROEs reported represent true proximity between protons. ROESY experiments were carried out at different mixing times (100, 200 and 300 ms), with the transmitter set in the middle of the NH region, and at a field strength of 2.0 kHz. This technique has been used for peptides from four to 27 amino-acids, and the resulting spectra were basically free of any cross-peaks that would adversely influence the interpretation[12,22,29,30].

Discussion

The backbone conformations of the tetrapeptides are defined by the torsion angles ϕ, ψ and ω of each residue, which, in principle, can be derived from the Hα-NH coupling constants and the strength of the nuclear Overhauser enhancements[31-35]. The conformation of side-chains is analyzed in terms of side-chain rotamer populations[36] for which clearly resolved Hα-Hβ and Hα-Hβ' coupling constants are a prerequisite. This analysis is complicated when the studied peptide does not possess a well-defined or rigid conformation. In the case of tetramers, the peptides possess an unlimited number of conformations

and all measurable parameters are therefore population-weighted averages of the conformational states available to the peptide, which precludes a numerically precise interpretation of the data. A list of accepted criteria for turn formation, some of which are considered more definitive than others, can be used[5], which includes: (a) $d_{NN}(i, i+1)$ NOE connectivities between residues 3 and 4 of the turn (type I turns also have a $d_{NN}(2,3)$ connectivity); (b) a $d_{\alpha N}(2,4)$ NOE connectivity; (c) appropriate $^3J_{HN\alpha}$ coupling constants (not always seen); (d) a low temperature coefficient for the NH of residue 4 of the turn, indicating hydrogen bonding (not always seen). Such a hydrogen bond is not a requirement for turn formation, at least not in proteins[37]; (e) characteristic CD or vibrational spectra.

In our previous study of DEKS[11], four pieces of evidence were clearly observed (Table 4.3), i.e. the $^3J_{HN\alpha}$ coupling constant for E^2 was rather low (4.8 Hz), the temperature dependence for S^4 was quite low (2.3 ppb/K), and NOE connectivities between NH(E^2) and NH(K^3) (weak), NH(K^3) and NH(S^4) (medium) were observed. A CD spectrum showed that DEKS did possess a proportion of ordered structure. Based on this evidence, we concluded that DEKS formed a type I β-turn in the solution. In the present study, basically the same set of data were observed for DERS and NEKS, i.e. a small coupling constants for E^2 (5.0 Hz) in both cases, a lower temperature dependence for S^4 (2.7 ppb/K for DERS and 3.8 ppb/K for NEKS), ROEs between NH(E^2) and NH(R^3) (weak), NH(R^3) and NH(S^4) (medium) for DERS, NH(K^3) and NH(S^4) (weak) for NEKS. In view of the NMR data on these two molecules, which are summarized in Table 4.3, we believe that NEKS and DERS form a type I β-turn in solution, as does DEKS. It must be noted that one piece of evidence leading to the detection of a β-turn, the $d_{\alpha N}(2,4)$ ROE connectivity, was missing in all three molecules. The reason for this might be that the population of β-turn in the

solution is so low that the experimental method is not sufficiently sensitive for this connectivity to be detected, since, in a type I β-turn structure, the $d_{\alpha N}(2,4)$ ROE is a much weaker interaction (3.6 Å) than the two d_{NN} ROEs (2.6 and 2.4 Å)[24]. The detection of a sequential $d_{\alpha N}(2,3)$ ROE, which is much stronger than the $d_{NN}(2,3)$ or $d_{NN}(3,4)$ ROEs, shows that in all three peptides, a smaller proportion of molecules in the type I β-turn conformation coexist with a larger proportion in extended conformations. A strong $d_{\alpha N}(2,3)$ ROE is only consistent with an extended conformation (2.3 Å) but cannot arise from the β-turn structure (3.4 Å). It has to be noted that in the case of small peptides in solution, it is impossible to classify all the conformations into two classes only, for example, β-turn and extended. This makes it impossible to assess the ratio between the molecules in the extended conformation and those adopting a β-turn in a numerically precise manner, based only on the NH-Hα coupling constants, the temperature coefficients and ROE intensities. The lack of standard values for the temperature coefficient of each residue in a certain conformation make it even worse. Nevertheless, based on the present NMR data, we conclude that approximately the same proportion of type I β-turn exists in solutions of DEKS and DERS. NEKS has a smaller population of β-turn, since the $d_{NN}(2,3)$ ROE, which is weaker than the $d_{NN}(3,4)$, was almost invisible in the ROESY spectrum and the temperature coefficient of S^4 was relative large.

Apparently the charge-charge interaction is dominant with regard to the secondary structure, especially in low dielectric constant solvents such as the methanol/water mixture[22]. Salt-bridging has been observed previously in model tetrapeptides[38] where the charge interaction was observed between moieties 1 and 4. Recently salt-bridging has been observed to be a stabilizing force in the formation of an α-helix[8], where the salt-bridge sets the amino terminal boundary of the helix. Previously, we postulated[11] that a salt-bridge between the side-

chains of E and K in DEKS, stabilized the β-turn conformation. For DERS where K is replaced by R, the β-turn is still present. By replacing E in DEKS or DERS by Q, the β-turn disappeared in DQKS and DQRS. Under low pH conditions, all evidence for β-turns disappeared for the three molecules. In the case of NEKS, this is significant as there is only one possible charge-charge interaction. These results clearly showed the importance of the salt-bridge in stabilizing the β-turn. Another interesting result is that the charge on the side-chain of the first-residue also stabilized the β-turn conformation. Replacing D in DEKS by N still maintains the β-turn in NEKS, but the population of the β-turn drops. Replacing D in DERS by N destroys the β-turn in NERS. All other replacements tested also destroyed the β-turn. It seems that a minus/minus/plus charge pattern and the exact lengths of the first and second side-chains are both important. These charges set up such a delicate balance to stabilize a β-turn that the loss or addition of a side-chain methylene group (D^1 to E^1 or E^2 to D^2) is sufficient to disrupt the β-turn conformation as in the case reported earlier[39], where an E to D substitution disrupted the balance between several charge-charge interactions, resulting in partial unfolding of an enzyme.

It is of interest to compare our results with the secondary structure predictions derived using the parameters of Wilmot and Thornton[13] (Table 4.4). Of the sixteen tetrapeptides studied, only five were not predicted to adopt a β-turn. Three of the five carried Q in position 2 and therefore could not form the salt-bridge. In the other two, D at position 1 was replaced by E. A general observation is that this method leads to overprediction of β-turns in the isolated tetrapeptides studied here. The three tetrapeptides for which we did find evidence by NMR of β-turn formation (DEKS, NEKS and DERS) were all predicted to adopt a type I β-turn but were by no means those assigned the highest probabilities for doing so. In fact the prediction for DEKS was the

weakest of all and DERS and NEKS were ranked seventh and ninth out of eleven, respectively. Even more striking was the total lack of evidence for β-turn formation in NPDM, NSDM and NDDS, the three tetrapeptides with by far the strongest predictions. It could be argued that the use of parameters derived from a protein structure data base for prediction of β-turns in small peptides in solution is inappropriate. However the β-turn is generally considered to be a unit of secondary structure which derives its stability from local steric factors and hydrogen bonding. Consequently we would have anticipated being able to detect at least some evidence for β-turns under the structure-promoting, low dielectric constant conditions used here. These findings could be reconciled if, in proteins the β-turns fall into two classes - those that are intrinsically stable and those that depend on long-range interactions (for example between strands of β-sheets) for their stability. Validation of this suggestion is clearly outside the scope of the present study.

To study this selectivity further, a series of calculations were carried out using the software "Biograf"[40]. In the program, the AMBER force field[41] was used. A distance-dependent dielectric constant was used to qualitatively simulate the fact that the peptides are in a methanol/water solvent and the intramolecular electrostatic interactions should die off more rapidly with distance than in the gas phase. But for the system of tetramers, as was expected, a simple energy minimization process could not provide any meaningful results to explain the experimental results, since the molecules possess too many conformations, undergoing rapid interchange between different conformers and have many degree of freedom and local energy minima.

References

1. Jaenicke, R. (1991), *Biochemistry* **30**, 3147-3161.
2. Dyson, H. J., Rance, M., Houghten, R. A., Lerner, R. A. & Wright, P. E. (1988), *J. Mol. Biol.* **201**, 161-200.
3. Dyson, H. J., Rance, M., Houghten, R. A., Lerner, R. A. & Wright, P. E. (1988), *J. Mol. Biol.* **201**, 201-217.
4. Wright, P. E., Dyson, H. J. & Lerner, R. A. (1988), *Biochemistry* **27**, 7167-7175.
5. Dyson, H. J. & Wright, P. E. (1991), *Annu. Rev. Biophys. Biophys. Chem.* **20**, 519-538.
6. Lewis, P. N., Momany, F. A. & Scheraga, H. A. (1971), *Proc. Nat. Acad. Sci. USA* **68**, 2293-2297.
7. Zimmerman, S. S. & Scheraga, H. A. (1977), *Proc. Nat. Acad. Sci. USA* **74**, 4126-4129.
8. Shoemaker, K. S., Kim, P. S., York, E. Y., Stewart, J. M. & Baldwin, R. L. (1987), *Nature* **326**, 563-567.
9. Smith, J. A. & Pease, L. G. (1980), *CRC Crit. Rev. Biochem.* **8**, 315-399.
10. Rose, G. D., Gierasch, L. M. & Smith, J. A.(1985), *Advan. Protein Chem.* **37**, 1-109.
11. Otter, A., Scott, P. G., Liu, X. & Kotovych, G. (1989), *J. Biomol. Struct. Dyn.* **7**, 455-476.
12. Otter, A., Kotovych, G. & Scott, P. G. (1989), *Biochemistry* **28**, 8003-8010.
13. Wilmot, C. M. & Thornton, J. M.(1988), *J. Mol. Biol.* **203**, 221-323.
14. Aue, A. P., Bartholdi, E. & Ernst, R. R.(1976), *J. Chem. Phys.* **64**, 2229-2246.
15. Bothner-By, A. A., Stephens, R. L., Lee, J., Warren, C. D. & Jeanloz, R. W.(1984), *J. Am. Chem. Soc.* **106**, 811-813.

16. Bax, A. & Davis, D. G. (1985), *J. Magn. Reson.* **63**, 207-213.

17. Marion, D. & Wüthrich, K. (1983), *Biochem. Biophys. Res. Commun.* **113**, 967-974.

18. Wider, G., Hosur, R. V. & Wüthrich, K. (1983), *J, Magn. Reson.* **52**, 130-135.

19. Mehlkopf, A. T., Korbee, D., Tiggelmann, T. A. & Freeman, R. (1984), *J. Magn. Reson.* **58**, 315-323.

20. Otting, G., Widmer, H., Wagner, G. & Wüthrich, K. (1986), *J. Magn. Reson.* **66**, 187-193.

21. Kessler, H., Griesinger, Ch., Kerssebaum, R., Wagner, R. & Ernst, R. R. (1987), *J. Am. Chem. Soc.* **109**, 607-609.

22. Otter, A., Scott, P. G., Cann, J. R., Vavrek, R. J., Stewart, J. M. & Kotovych, G. (1988), *J. Biomol. Struct. Dyn.* **6**, 609-625.

23. Griesinger, Ch. & Ernst, R. R. (1987), *J. Magn. Reson.* **75**, 261-271.

24. Wüthrich, K. (1986), *NMR of Proteins and Nucleic Acids*, Wiley-Interscience, New York.

25. Marion, D. (1985), *FEBS Letters* **192**, 99-103.

26. Neuhaus, D. & Keeler, K. (1986), *J. Magn. Reson.* **68**, 568-574.

27. Wagner, G. & Wüthrich, K. (1982), *J. Mol. Biol.* **155**, 347-366.

28. Noggle, J. H., & Schirmer, R. E. (1971), *The nuclear Overhauser effect*, Academic Press, New York.

29. Otter, A. & Kotovych, G. (1988), *Can. J. Chem.* **66**, 1814-1820.

30. Otter, A., Scott, P. G. & Kotovych, G. (1988), *Biochemistry* **27**, 3560-3567.

31. Karplus, M. J. (1963), *J. Am. Chem. Soc.* **85**, 2870-2871.

32. Pardi, A., Billeter, M. & Wüthrich, K. (1984), *J. Mol. Biol.* **180**, 741-751.

33. Richardson, J. S. (1981), *Advan. Protein Chem.* **34**, 167-339.

34. Leach, S. J., Némethy, G. & Scheraga, H. A. (1977), *Biochem. Biophys.*

Res. Commun. **75**, 207-215.

35. Bystrov, V. F. (1976), *Prog. Nucl. Mag. Reson. Spectrosc.* **10**, 41-81.

36. Jardetzky, O. & Roberts, G. C. K. (1981), *NMR in Molecular Biology*, Academic Press, New York.

37. Chou, P. Y. & Fasman, G. D. (1977), *J. Mol. Biol.* **115**, 135-175.

38. Sahal, D. & Balaram, P. (1986), *Biochemistry* **25**, 6004-6013.

39. Daar, I. O., Artymiuk, P. J., Phillips, D. C. & Maquat, L. E. (1986), *Proc. Natl. Acad. Sci. USA* **83**, 7903-7907.

40. BIOGRAF (Version 2.20), BioDesign, Inc.,199 S. Los Robles Ave., Pasadena, CA 91101.

41. Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta, S. Jr. & Weiner, P. (1984), *J. Am. Chem. Soc.* **106**, 765-784.

Table 4.1 Summary of Experimental Parameters Used in the Two-Dimensional NMR Experiments[a]

Parameters	COSYPH	ROESYPH[b]
Sweep width in F2 (Hz)	3600	3600
Sweep width in F1 (Hz)	1800x2	1800x2
Matrix size (F1xF2) before zero-filling	256x1K	256x1K
Matrix size (F1xF2) after zero-filling	512x2K	512x2K
Evolution time		
Initial value (us)	1	1
Increment (us)	139	139
Number of scans	48	64
(dummy scans)	(2)	(2)
Acquisition time (s)	0.57	0.57
Relaxation delay (s)	1.8	1.8
Other delay (ms)	--	300[c]
Window functions for 2D FT (F1/F2)[d]	S/S	S/S
Shifts of window function in fractions of π (F1/F2)	6/8	3/4

a) Some parameters are slightly different, depending on the molecule.

b) The ROESY experiment was performed several times with different transmitter frequencies and mixing times (see text).

c) Spin locking time at an average field strength of 4.0 kHz

d) S stands for a sine bell.

Table 4.2 Proton Chemical Shifts[a], Coupling Constants[b] and NH Temperature Coefficients[c] of the Peptides Studied[d]

Pep. Res.	NH	α	β[f]	γ[f]	δ[f]	ε	J(Hz)	ΔT (ppb/K)
NEKS								
NAc		2.03 (2.01)[e]						
Asn1	8.39 (8.35)	4.66 (4.66)	2.73,2.78 (2.69,2.80)	7.66,6.92 (7.64,6.92) (NH_2)[g]			7.2 (7.0)	6.4
Glu2	9.07 (8.46)	4.22 (4.31)	2.03 (1.94,2.10)	2.28 (2.46)			5.0 (6.6)	4.2
Lys3	8.27 (8.31)	4.33 (4.33)	1.82,1.88 (1.78)	1.47 (1.46)	1.68 (1.69)	2.97 (2.98)	7.3 (7.0)	5.1
Ser4	7.96 (8.017)	4.35 (4.38)	3.83,3.88 (3.81)				7.3 (7.4)	3.8
NH_2	7.45,7.13 (7.52,7.13)							

EEKS

NAc		2.03 (2.02)[e]						
Glu1	8.41 (8.30)	4.22 (4.29)	1.91,2.02 (1.95,2.09)	2.25 (2.44)			6.0 (7.0)	5.6
Glu2	8.78 (8.44)	4.26 (4.34)	1.91,2.02 (1.95,2.09)	2.25 (2.44)			6.5 (7.0)	5.3
Lys3	8.32 (8.33)	4.35 (4.34)	1.78,1.88 (1.75,1.86)	1.43 (1.44)	1.68 (1.68)	2.98 (2.97)	7.0 (7.5)	6.4
Ser4	8.11 (8.13)	4.36 (4.39)	3.83,3.88 (3.82,3.85)				7.5 (7.5)	5.3
NH_2	7.13,7.52 (7.12,7.59)							

DDKS								
NAc		2.02 (2.02)[e]						
Asp1	8.25 (8.46)	4.58 (4.66)	2.60,2.70 (2.73,2.79)				7.0 (7.5)	6.3
Asp2	8.49 (8.50)	4.51 (4.66)	2.68,2.73 (2.86,2.91)				6.9 (8.0)	3.3
Lys3	8.35 (8.16)	4.31 (4.33)	1.80,1.87 (1.79,1.86)	1.46 (1.45)	1.65 (1.68)	2.98 (2.97)	7.6 (7.5)	5.6
Ser4	8.16 (8.01)	4.30 (4.38)	3.84,3.90 (3.84)				7.2 (7.5)	4.1
NH_2	7.44,7.11 (7.48,7.12)							

DQKS

NAc		2.07						
Asp1	8.23	4.55	2.62,2.68				7.0	6.7
Gln2	8.56	4.24	1.97,2.15	2.36	7.59,6.84 (NH_2)		6.5	7.3
Lys3	8.40	4.29	1.84	1.43	1.67	2.97	7.0	5.4
Ser4	8.05	4.36	3.83,3.87				7.0	4.2
NH_2	7.46,7.15							

NQKS

NAc		2.03						
Asn1	8.37	4.65	2.70,2.79	7.64,6.94 (NH_2)			7.5	6.2
Gln2	8.53	4.38	2.03,2.11	2.35	7.58,6.86 (NH_2)		7.0	6.3
Lys3	8.36	4.32	1.79,1.88	1.42	1.68	2.97	7.0	5.6
Ser4	8.09	4.38	3.84				7.0	5.6
NH_2	7.55,7.15							

DERS								
NAc		2.01 (2.02)[e]						
Asp1	8.19 (8.44)	4.61 (4.67)	2.70,2.74 (2.76,2.89)				7.5 (7.0)	6.9
Glu2	9.16 (8.44)	4.18 (4.33)	2.20 (1.98,2.14)	2.30 (2.44)			5.0 (7.0)	4.8
Arg3	8.50 (8.24)	4.24 (4.35)	1.90 (1.76,1.85)	1.69 (1.65)	3.20 (3.20)	(7.37) (NH)	7.0 (7.1)	3.9
Ser4	7.98 (8.04)	4.34 (4.39)	3.88,3.92 (3.83)				7.3 (7.5)	2.7
NH_2	7.31,7.16 (7.55,7.12)							

NERS

NAc		2.03						
Asn1	8.39	4.66	2.71,2.81	7.65,6.92 (NH_2)			7.0	6.8
Glu2	8.89	4.23	2.03	2.28			6.0	5.1
Arg3	8.31	4.35	1.78	1.64	3.20		7.5	5.6
Ser4	8.03	4.36	3.83,3.87				7.5	4.7
NH_2	7.51,7.15							

EERS

NAc		2.02						
Glu1	8.42	4.22	2.25	1.90			6.5	5.6
Glu2	8.82	4.25	2.25	2.02			6.0	5.3
Arg3	8.36	4.36	1.78	1.64	3.20	7.54 (NH)	7.2	6.3
Ser4	8.14	4.37	3.84,3.88				7.5	5.2
NH_2	7.42,7.15							

DDRS

NAc		2.01 (2.01)[e]						
Asp1	8.24 (8.45)	4.59 (4.66)	2.61,2.72 (2.81,2.87)				7.1 (7.1)	6.4
Asp2	8.49 (8.52)	4.53 (4.67)	2.70 (2.75,2.91)				6.9 (7.5)	4.2
Arg3	8.39 (8.18)	4.28 (4.35)	1.84,1.91 (1.74,1.86)	1.67 (1.65)	3.20 (3.20)	7.67 (7.32) (NH)	7.2 (7.5)	4.8
Ser4	8.17 (8.07)	4.30 (4.38)	3.87,3.91 (3.84)				7.2 (7.5)	3.9
NH_2	7.44,7.13 (7.51,7.13)							

NDRS

NAc		2.03						
Asn1	8.34	4.70	2.71,2.85	7.62,6.89 (NH_2)			7.5	6.1
Asp2	8.54	4.51	2.64,2.69				6.8	4.6
Arg3	8.23	4.32	1.80,1.86	1.67	3.19		7.5	5.2
Ser4	8.26	4.62	4.29				7.5	6.4
NH_2	7.47,7.21							

DQRS

NAc	2.00							
Asp1	8.24	4.57	2.64,2.73				7.0	7.3
Gln2	8.66	4.25	1.98,2.19	2.38	7.57,6.83 (NH_2)		7.0	8.3
Arg3	8.45	4.25	1.87	1.67	3.19	7.71 (NH)	7.0	4.5
Ser4	8.04	4.36	3.83,3.87				7.2	4.5
NH_2	7.41,7.15							

DKES

NAc		2.01 (2.01)[e]						
Asp1	8.23 (8.43)	4.57 (4.66)	2.63,2.70 (2.79,2.91)				6.9 (7.1)	5.6
Lys2	8.46 (8.43)	4.25 (4.29)	1.85 (1.72,1.82)	1.47 (1.43)	1.68 (1.68)	2.99 (2.97)	6.1 (7.4)	6.9
Glu3	8.57 (8.26)	4.25 (4.36)	2.05 (2.45)	2.27 (2.00,2.14)			6.5 (7.4)	3.3
Ser4	8.08 (8.01)	4.35 (4.37)	3.82,3.90 (3.80,3.86)				7.5 (7.8)	3.3
NH_2	7.42,7.12 (7.50,7.13)							

NPDM

NAc		1.98					
Asn1	8.31[9]	4.95	2.65,2.84	7.69,7.07 (NH_2)		7.5	7.2
Pro2		4.38	2.17,2.28	2.00	3.78 3.87		
Asp3	8.25	4.53	2.62,2.68			7.0	4.1
Met4	8.00	4.41	1.91,2.00	2.46,2.57	2.09	8.0	4.5
NH_2	7.49,7.09						

NSDM

NAc		2.02					
Asn1	8.32	4.73	2.68,2.82	7.65,6.95 (NH_2)		7.5	6.3
Ser2	8.40	4.41	3.81,3.89			7.5	6.7
Asp3	8.38	4.57	2.62,2.75			6.0	4.4
Met4	8.14	4.38	1.97,2.16	2.47,2.59	2.09	8.0	4.5
NH_2	7.53,7.03						

NDDS						
NAc		2.01				
Asn1	8.34	4.70	2.72,2.79	7.65,6.92 (NH_2)	7.5	6.4
Asp2	8.38	4.59	2.69		7.5	5.5
Asp3	8.34	4.60	2.69		7.5	4.3
Ser4	8.25	4.29	3.89		7.5	3.9
NH_2	7.54,7.10					

a) Chemical shifts are reported in ppm relative to the methyl group of CD_3OH (3.30 ppm downfield from TMS).

b) Coupling constants in Hz refer to $J_{NH\text{-}H\alpha}$ and were taken from 1D spectra.

c) The effect of temperature on the chemical shifts of the exchangeable protons is an average value of temperature coefficients measured at either eight or four points between 284K and 300K, at 400 MHz as discussed in the text.

d) The spectra were measured over a pH range of 5.5 to 6.6 which is well above the pKa's of D and E with the exception of NQKS (pH 5.3). The concentrations studied at this pH were between 3.9 and 6.8 mmol/L in 60% CD_3OH / 40% H_2O at ambient temperature. Six tetrapeptides were also measured at low pH, ranging from 1.9 to 3.2, well below the pKa's of D and E. Higher concentrations (6.8 to 23.1 mmol/L) were used to confirm the absence of even small ROE's under these conditions.

e) The numbers in brackets are for the samples at low pH.

f) The pairs of protons x (x=α, β, γ ...) could not be assigned stereospecifically.

g) The chemical shift for the cis-isomer is 8.402 ppm.

Table 4.3 **Summary of ^1H Chemical Shifts[a], Coupling Constants and NH Temperature Coefficients of DEKS, DERS and NEKS in CD_3OH/H_2O (60/40)**

DEKS[b]				
	NH [ppm]	$J_{NH-\alpha}$ [Hz]	NH f(T) [ppb/K]	ROE (NH-NH)
D	8.25	7.4	6.9	
	(8.34)	(7.3)	(4.9)	
E	9.17	4.8	3.7	
	(8.61)	(6.5)	(6.7)	weak
K	8.43	7.3	3.8	
	(8.37)	(7.2)	(6.5)	medium
S	7.97	7.3	2.3	
	(8.04)	(7.4)	(4.2)	

DERS				
	NH [ppm]	$J_{NH-\alpha}$ [Hz]	NH f(T) [ppb/K]	ROE (NH-NH)
D	8.19	7.5	6.9	
	(8.44)	(7.0)		
E	9.16	5.0	4.8	
	(8.44)	(7.0)		weak
R	8.50	7.0	3.9	
	(8.24)	(7.1)		medium
S	7.98	7.3	2.7	
	(8.04)	(7.5)		

NEKS				
	NH [ppm]	$J_{NH-\alpha}$ [Hz]	NH f(T) [ppb/K]	ROE (NH-NH)
N	8.39	7.2	6.4	
	(8.35)	(7.0)		
E	9.07	5.0	4.2	
	(8.46)	(6.6)		
K	8.27	7.3	5.1	
	(8.31)	(7.0)		weak
S	7.96	7.3	3.8	
	(8.02)	(7.4)		

a) Chemical shifts in ppm relative to the methyl signal of CD_3OH (3.30 ppm downfield from TMS).

b) These data on DEKS were published previously[11] at 283K, pH 6.4 at a concentration of 10 mM. The values in brackets were obtained at pH 3.5, otherwise the same experimental conditions were used.

Table 4.4 Probabilities of β-turn Formation in Tetrapeptides[a]

Peptide	Prediction	
DEKS	Type I	1.012[b]
DQKS	----[c]	<1.0
DDKS	Type I	1.094
NEKS	Type I	1.361
NQKS	----[c]	<1.0
EEKS	----[c]	<1.0
DKES	Type N	1.845[d]
DERS	Type I	2.228[e]
DQRS	----[c]	<1.0
DDRS	Type I	2.410[e]
NDRS	Type I	3.241[e]
EERS	----[c]	<1.0
NERS	Type I	2.997
NPDM	Type I	23.327
NSDM	Type I	17.244
NDDS	Type I	4.784[e]

a) see reference 13.

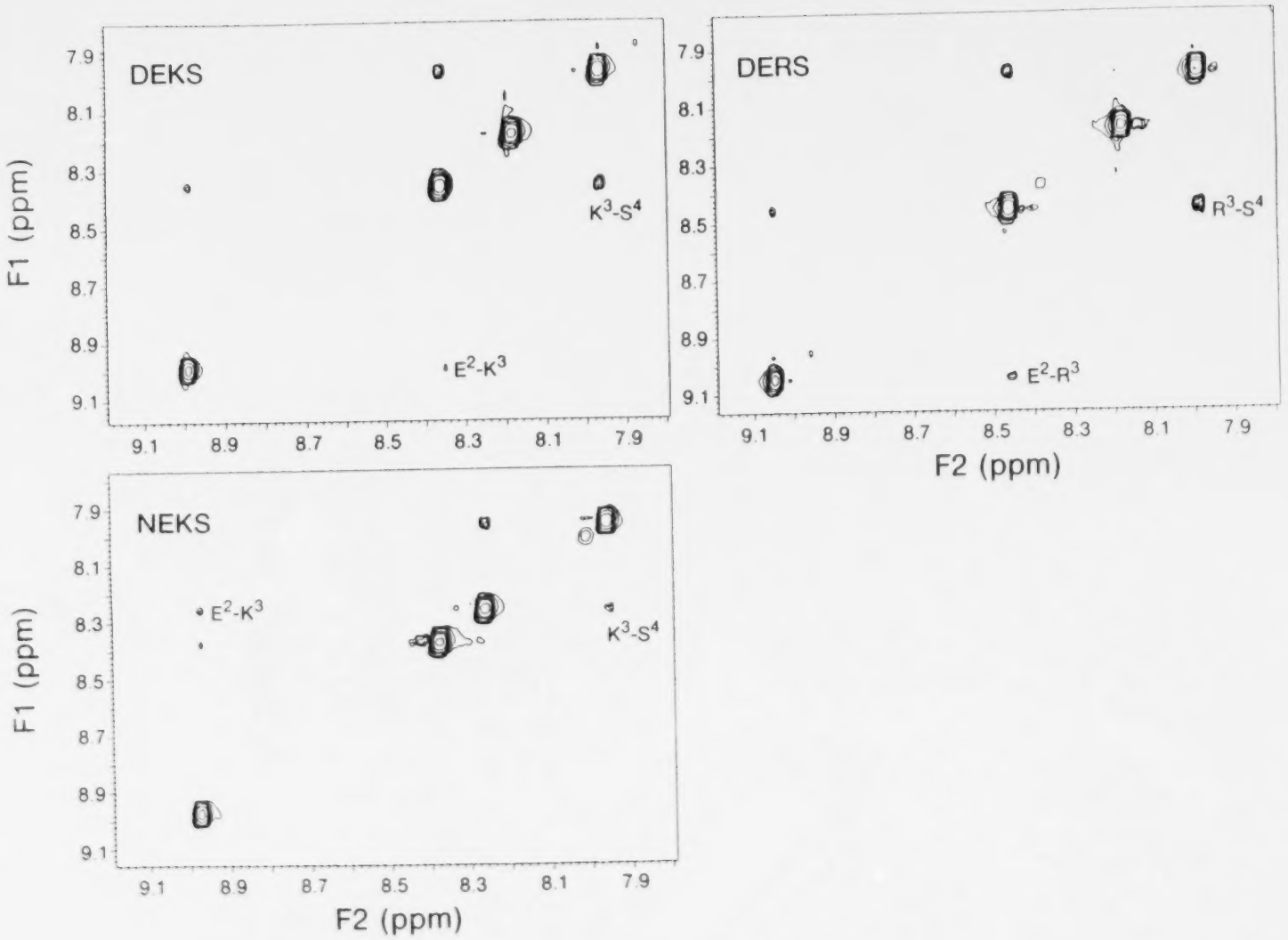
b) Probability in multiples of cut-off probability.

c) Below cut-off.

d) Type I also possible but at lower probability.

e) Type N also possible but at lower probability.

Figure 4.1 NH-NH region of the 500 MHz phase-sensitive ROESY experiments on DEKS, DERS and NEKS in CD_3OH/H_2O (60/40). The spectra were recorded at 20°C. The sample concentration was 5 mM, pH 5.8. The spin-locking time was 300 ms at a field strength of 5.0 kHz. The cross-peaks are of opposite sign with respect to the diagonal.


DEKS
DERS
NEKS
F1 (ppm)
F2 (ppm)
7.9
8.1
8.3
8.5
8.7
8.9
9.1
K^3-S^4
R^3-S^4
E^2-K^3
E^2-R^3

CHAPTER 5

Proton Magnetic Resonance Studies of Bradykinin Antagonists*

Introduction

Bradykinin (BK) is a linear peptide hormone (Arg^1-Pro^2-Pro^3-Gly^4-Phe^5-Ser^6-Pro^7-Phe^8-Arg^9) and has been implicated in a multitude of pathophysiological processes[1,2]. Of particular significance is its role as a potent pain-producing agent. Recently, it has been suggested that bradykinin may be associated with the symptoms of the common cold[3,4]. Because of these physiological functions, a bradykinin receptor antagonist may have significant therapeutic value, and the synthesis of these antagonists has been the subject of intensive research.

Many bradykinin antagonists have been synthesized with the aim of increasing the potency, selectivity and lifetime of these antagonists[5,6]. In almost all of these syntheses, analogs of natural amino acids were used to impose specific conformational restraints[7,8] in order to obtain insight about their bioactive conformation. So far, BK analogs having bulky, beta-branched or cyclic D-aliphatic residues at position 7 combined with bulky or cyclic L-aliphatic residues at position 8 have yielded the best antagonists[9,10], as have bradykinin antagonist dimers[11].

Earlier conformational studies in solution[12,13] on bradykinin itself and an antagonist led to a hypothesis that the difference between an agonist and an

---

* A version of this chapter has been accepted by *Biopolymers*. Liu, X., Stewart, J. M., Gera, L., & Kotovych, G. (1993), *Biopolymers*, in press.

antagonist is related to the type of β-turn adopted at the C-terminus together with the required orientation of the side chains[12]. To challenge this hypothesis, Kyle et al.[14,15] designed and prepared several constrained bradykinin analogs, which were assumed to have an inherently stabilized β-turn geometry at their C-terminus based on conformational analysis using empirical calculations. Two of them were found to be competitive antagonists with big bulky side-chains at position 7 or 8, such as DArg^0-Arg^1-Pro^2-Hyp^3-Gly^4-Thi^5-Ser^6-DTic^7-Tic^8-Arg^9 (Tic = tetrahydroisoquinoline carboxylic acid residue), but NMR data were not reported for them. More recently, two bradykinin analogs containing α-MePro at position 3 or position 7 were reported to show reversed-turn conformations at both Pro^2-Phe^5 and Ser^6-Arg^9, respectively, in aqueous solution by NMR, but no activity data were reported on these molecules[16].

In the present study, we investigated by ^{1}H NMR the solution conformation of [DArg^0, Hyp^3, Thi^5, DCpg^7, Cpg^8]-BK (I) (Cpg = α-cyclopentyl-glycine; Hyp = *trans*-4-hydroxy-L-proline, Thi = β-(2-thienyl)-L-alanine) (Figure 5.1), an extremely potent bradykinin antagonist and several related antagonists, which differ from each other only by one or two residues, in order to provide more experimental evidence for the hypothesis that a β-turn in the four C-terminal amino acid residues of bradykinin analogs might be a prerequisite for their high activities[12]. This study was carried out on the antagonists in the free state as it is impossible to carry out NMR studies when they are bound to the receptor. These antagonists function at nanomolar concentration, far below the limit of NMR detection sensitivity. The solvent for this study is a methanol/water (80/20 v/v) mixture, which is a structure-inducing solvent that has been used in the study of other small peptides[17]. The sequences for all the peptides studied are listed in Table 5.1 and pharmacological data are listed in Table 5.2. This ^{1}H NMR study is important to help clarify the role each residue plays and to

understand the geometric requirements for bradykinin antagonists. In a separate paper, we present a similar NMR study on two agonist-antagonist pairs[18].

Materials and Methods

Materials and Sample Preparation

The peptides were synthesized, purified and assayed in the classical BK assay systems as reported previously[6]. The samples used in all ^1H NMR experiments were prepared by dissolving purified solid peptides in 160 μl H_2O. The pH was then adjusted by adding small aliquots of diluted NaOH or HCl, followed by adding 640 μl CD_3OH to make the solvent an 80/20 methanol/water mixture. Before the NMR experiments, the samples were degassed with argon. The sample concentrations were between 0.9 and 1.4 mmol/L and the pH of the samples was in the range of 5.2 to 5.4. Aggregation is minimized by these low concentrations, together with the fact that the active peptides contain four positive charges. Peptide I was diluted by a factor of ten with no substantial changes in chemical shifts, also indicating the absence of aggregation.

NMR Spectroscopy

All experiments were carried out on a Varian UNITY 500 NMR spectrometer using a proton-selective 5 mm probe with a 90° proton pulse length of 8.8 to 10.0 μs at normal power levels. During the experiments the temperature was always controlled to ±0.1°C and set to 0°C except for the temperature studies. For FID acquisitions and data processing, VNMR 3.1 and 3.2 software were used on a SUN 4/330 and a SUN 4/470 workstation, respectively. The proton chemical shifts were measured relative to the undeuterated fraction of the methyl group of CD_3OH (Merck, Sharp & Dome,

MD-67, 99.9% D) at 3.30 ppm. All 2D spectra were acquired non-spinning, whereas for 1D experiments the spinner was turned on.

In all NMR experiments, the water resonance from the solvent was suppressed during the relaxation delay by presaturation. The decoupler offset was set equal to the transmitter offset to prevent phase distortions around the water peak due to imperfect cancellation of the dispersive component of residual water magnetization[19]. Before each experiment, the delays before and after the reading pulse (alfa delay) were carefully adjusted to give a flat and undistorted baseline[20].

One-dimensional proton spectra were recorded with a sweep width of 5000Hz, a relaxation time of 2s and 16K (zero-filled to 32K) data points. Usually a total of 32 scans were accumulated, and the resulting FID's were resolution enhanced by means of a Gaussian multiplication with a line broadening factor of typically -3 Hz.

The temperature dependence of the amide protons was calculated by measuring the chemical shifts at -5, 0, 5, 10, 15, 20 and 25 °C for all peptides and then analyzed by means of a linear regression. The controller temperatures were calibrated using a methanol sample provided by Varian[21] and were found to be in the range of +/-1.0 °C of the predicted temperature.

Two-dimensional experiments were measured with 2048 (zero filled to 4096) or 4096 data points along t_2 and 256 t_1 increments. Spectral widths of 5000 Hz in both dimensions and a relaxation delay of 1.5 to 2.0 s were employed. All two-dimensional experiments were acquired in the phase-sensitive mode by using the hypercomplex technique known as the States-Haberkorn-Ruben method[22]. The TOCSY[23] experiments were carried out by using the basic pulse sequence proposed by Bax[24]. A 2ms trim pulse preceded the 70-80ms MLEV-17 spin-lock (field strength 7.0 kHz). 16 to 32 scans were

accumulated per experiment. A $\pi/2$ shifted, squared sine bell weighting function was used in both dimensions. The ROESY[25,26] experiments were recorded with the pulse sequence 90°-t1-90°-SL-90°-FID. Spin-lock conditions (SL) were achieved by a series of hard pulses of 30° flip angle[27] and the 90° pulses immediately preceding and following the spin locking period were added to compensate for the non-negligible offset dependence of the cross-peaks (compensated ROESY)[28]. The experiments were carried out with mixing times of 100 to 200 ms, spin lock field strength of 5.0 kHz and 32 scans per FID. In all cases, $\pi/2$ shifted, squared sine bell weighting functions were used. The NOESY experiments[29,30] were carried out with the 90°-t1-90°-mix-90°-FID pulse sequence, 200 or 300 ms mixing time and 32 scans per t1 increment. Water suppression was achieved by decoupler saturation at all times except during the acquisition period. A 10% random delay was incorporated in the mixing delay to suppress zero quantum coherences. For TOCSY, ROESY and NOESY experiments, the compensation for first point distortion caused by analog filters[31] was achieved by reducing the first point of each FID and each t1-interferogram by empirically determined values[20]. For all the 2D spectra, baseline corrections were not necessary.

Results

Chemical Shift Assignments and Coupling Constants

The phase-sensitive TOCSY spectra of the six peptides in methanol/water (80/20) alone allowed the identification of all unique amino acids in a sequence[23]. Then, a sequential analysis was performed using the NOESY or ROESY spectra[32] to achieve the complete assignment. These assignments rely on H_α-NH(i, i+1), NH-NH(i, i+1) and for Pro residues, H_α-$H_{\delta,\delta'}$(i, i+1) connectivities. The NH resonance of the first N-terminal residue is always

absent due to fast exchange with the solvent since the peptides are not protected. In addition, the NH resonances of Arg^1 for peptides I to III are usually very weak and can only be seen at low temperature. However, the spin systems for these Arg resonances can be easily identified through their NH resonance in the side chain. It was difficult to assign specifically each of the aliphatic ring protons in the Cpg residues because of spectral overlap. At best, the four protons at positions 2 and 5 can be assigned, but the protons at positions 3 & 4 in each ring cannot be resolved. The cis isomers due to the X-Pro or X-Hyp peptide bond are not observed. This observation is consistent with the suggestion based on NMR experiments run in a lipophilic environment that all amide bonds remain trans in both agonists and antagonists[15].

The assigned resonances for all the peptides are summarized in Table 5.3. As an example of the experimental data, the relevant section of the TOCSY contour plot for peptide I, recorded with a mixing time of 80 ms is shown in Figure 5.2, whereas the NH/NH section of its NOESY spectrum recorded with a mixing time of 300 ms is depicted in Figure 5.3.

Peptides I and II exhibit many similarities in terms of chemical shifts, which is understandable since the peptides only differ by one residue (Thi^5 in I is replaced by Phe^5 in II). From peptide I to peptide II, very little change in the chemical shifts of the amide protons was observed. In these two peptides, the NH resonances for Gly^4 are to low field and the two α-protons are well separated, indicating less flexibility in this region of the molecules. Similar behavior for Gly^4 was observed for peptide III. However, it is interesting to notice that the chemical shift for some NH protons changes when one or two residues in a peptide are changed. When $DCpg^7$ in peptide II is replaced by Cpg^7, resulting in peptide III, the chemical shifts for the NH protons of Cpg^8 shift to high field (8.47 to 8.17 from II to III). The chemical shifts of the other NH

resonances for the neighboring residues were also changed, but to a lesser extent.

All vicinal coupling constants of the type $^3J_{\alpha NH}$ could be measured in the one-dimensional spectra, at different temperatures as required to minimize spectral overlap. None of the coupling constants changed by more than 0.3 Hz over the temperature range -5 to 25°C, an indication that the conformation does not change significantly. All the observed $^3J_{\alpha NH}$ values are close to the time-averaged value of 6.3 Hz for free rotation about the N-C_α bond. The expected value for residues at positions i+1 and i+2 for β-turns in proteins are predicted to be approximately 4 and 9 Hz, respectively[32], but because of the presence of Pro^2 and Hyp^3 (or Pro^3), residues for which coupling constants are not available, no firm conclusions about preferred conformations can be drawn from these data.

Temperature Dependence of NH-chemical Shifts

An amide proton involved in a stable intramolecular hydrogen bond, or one inaccessible to solvent for steric reasons typically shows a reduced temperature coefficient in the range 0 to -3 ppb/K[33]. Although such low temperature coefficients do not necessarily mean the presence of a folded structure, they are important indicators of secondary structure when combined with NOE data, especially in the case of small peptides. The temperature coefficients for the amide protons of peptides **I** to **IV** are listed in Table 5.3. The important data for peptides **I** and **II** are summarized in Figure 5.4. In all cases, the chemical shifts of the amide protons varied linearly with temperature with correlation coefficients better than 0.988, indicating that major conformational changes are not present in the temperature range under study. In peptides **I** and **II**, which behave similarly in terms of temperature coefficients, Thi^5 and Phe^5

displayed low temperature coefficients of -2.4 and -2.9, respectively, which suggest that these protons might be involved in a turn. In comparing peptides **I** and **II** with peptide **III**, the low temperature coefficients at position 5 disappeared.

It is interesting to compare the temperature coefficients of Phe^5 or Thi^5 and Arg^9 in all the peptides studied, which are summarized in Table 5.2. A general trend can be seen, namely that the two active antagonists (**I** & **II**) have a low value at position 5 and high value at position 9. Less active antagonists have a lower value at position 9 and a higher one at position 5 (peptides **III**, **IV**). In the case of agonists, the NH of Arg^9 was involved in an H-bond and displayed the lowest temperature coefficients in the peptides[18].

NOE Information

Both ROESY and NOESY experiments were carried out for all the peptides. In all the NOESY experiments, the cross-peaks have the same sign with respect to the diagonal peak, which means that the NOEs are negative, probably because the experiments were carried out at 0°C. These two experiments showed similar results, but with different NOE intensities. For peptides **I** and **II**, the NOESY experiments showed stronger NOEs than the ROESY experiments, thus revealing more information. For peptide **IV**, NOESY and ROESY were about the same. For peptide **III**, the ROESY experiment appeared to be better than the NOESY. For peptide **I**, the ROESY experiments were carried out with different mixing times and different strengths of the spin-lock to detect any artifacts[34,35]. NOESY experiments were run with different mixing times as well. For peptides **I** and **II**, the cross-peak volumes of the NOESY at a mixing time of 300 ms were integrated with a correction for t1 or t2 ridges and for slight phase imperfection by measuring both cross-peaks. Where applicable, the volumes were normalized for the contributions of more than two

protons to the cross-peak by a factor of $(2N_aN_b)/(N_a+N_b)$, where N_a and N_b are the number of protons responsible for the diagonal peaks a and b, respectively[36]. The results based on these integrals for peptides **I** and **II** are summarized in Figure 5.4.

For each peptide, a set of expected, consecutive, strong $d_{\alpha N}(i, i+1)$ and weak $d_{NN}(i, i+1)$ were identified in the ROESY or the NOESY spectra. Numerous other ROEs or NOEs were observed between protons in neighboring residues. Figure 5.3 shows the NH/NH region of a 500 MHz phase-sensitive NOESY experiment of peptide **I** at a mixing time of 300ms. As can be seen, the NHs of residue 5 and 6 overlap at 0°C and have a similar temperature coefficient for peptides **I-III**, making it impossible to decide whether cross-peaks exist between those two NHs as they are too close to the diagonal.

In addition, some medium range NOEs were identified in peptides **I** and **II**. In peptide **II**, they were between the NH of Phe^5 and each of the following protons: $H_\alpha(Hyp^3)$, $H_\beta(Pro^2)$. In peptide **I**, they were between H_β of Pro^2 and NH of Thi^5. The NOE between H_α of Hyp^3 and NH of Thi^5 is not observable because of overlap of the H_α protons of Thi^5 and Hyp^3. These medium range NOEs are necessary to prove a turn-like structure. Extra NOE information cannot be extracted from the data on peptide **III** since the NHs of residues 5 to 8 overlap. For peptide **IV**, in addition to the usually observed intra- and inter-residue ROEs associated with an extended conformation, some unusual ROE patterns were observed around residues 7 and 8. A very strong NH-$H_\beta(Phe^8)$ cross-peak was found but no NH/$H_{\beta'}(Phe^8)$ and also no NH-$H_\alpha(Phe^8)$ interactions were evident. Furthermore, ROEs between $H\gamma(DCpg^7)$ at 0.72 ppm and NH(Phe^8) as well as between $H\gamma\gamma'(DCpg^7)$ at 0.72/0.91 ppm and the aromatic ring protons of Phe^8 were observed, indicating rigidity within this region of the molecule.

Discussion

The conformations of peptides are defined by a set of torsion angles for each residue which, in principle, can be derived from coupling constants, the strength of the nuclear Overhauser enhancements and amide proton exchange rate and/or temperature coefficients. Unfortunately, these general rules can only be applied when the studied peptide possesses a relatively well-defined or rigid conformation. In the case of small peptides like the bradykinin antagonists, the peptides possess a high number of conformations and all measurable parameters are therefore population-weighted averages of the conformational states available to the peptide, which precludes a numerically precise interpretation of the data. But a list of accepted criteria for turn formation, some of which are considered more definitive than others, can be used[37], which include: (a) $d_{NN}(3,4)$ NOE connectivities (type I turns also have a $d_{NN}(2,3)$ connectivity); (b) a $d_{\alpha N}(2,4)$ NOE connectivity; (c) appropriate $^{3}J_{HN\alpha}$ coupling constants (not always seen); (d) a low temperature coefficient for the NH of residue 4 of the turn, indicating hydrogen bonding. Although such a hydrogen bond is not a requirement for turn formation, at least not in proteins, it is a very important stabilizing factor in small peptides; and (e) characteristic CD or vibrational spectra. In light of these criteria, it is possible to draw only a qualitative conclusion about the presence of a turn in the peptides.

Two highly active bradykinin antagonists, peptides I & II, which differ from each other at position 5 (Thi^{5} vs. Phe^{5}), show almost identical NMR spectra in terms of chemical shifts, temperature coefficients and NOE information. Based on their sequential NOEs, extended conformations are dominant in solution for both molecules. A series of NH(i)-NH(i+1) NOEs, which were observed with a similar intensity, except between $DCpg^{7}$ and Cpg^{8}, are compatible with the

extended conformation. The relatively weaker NH/NH NOE between DCpg^7 and Cpg^8 may result partially from restrictions caused by the two bulky aliphatic side chains. The extended conformation for the C-terminal end is thus influenced by the bulky side-chains at DCpg^7 and Cpg^8. However, observed medium-range NOEs, i.e. H_α(Hyp^3) with NH(Phe^5), H_β(Pro^2) with NH(Phe^5) in peptide **II**, H_β(Pro^2) with NH(Thi^5) in peptide **I**, are not compatible with a fully extended conformation. These weak NOEs can only be explained by a reverse turn for this section of the peptides. This conclusion is supported by the low temperature coefficient of Phe^5 and Thi^5, respectively, indicating that their amide protons are involved in a hydrogen bond. Although one of the important criteria, namely the NOE between H_α(Hyp^3) and NH(Thi^5) in peptide **I** is unobservable because of the overlap of the Hα protons of Thi^5 and Hyp^3, this is observed in peptide **II** and hence can be inferred for peptide **I** because of the similarity between peptides **I** & **II**.

It is not surprising to find evidence for the turn-like structure between residues 2 and 5 in peptides **I** and **II**. According to Wilmot & Thornton's prediction[38], Pro in position (i+1) is the most strongly preferred amino acid for both type I & II β-turns, because the proline-inherent restriction of φ to approximately -60° satisfies the φ(i+1) requirement for both types. On the basis of the Chou and Fasman prediction[39], the sequence Pro^2-Pro^3-Gly^4-Phe^5 has a significant β-turn probability. Our results are consistent with a previous study of an antagonist[12] and of the α-methylproline analog[16] of bradykinin where a turn-like structure was detected from position 2 to 5. As in the previous studies[12,16], it is difficult to determine which type of turn is adopted by those four residues. The differences between the type I and the type II β-turns in terms of NMR information are a lack of NH(i+1)/NH(i+2) NOEs and stronger Hα(i+1)/NH(i+2) NOE in the type II β-turn[32]. Both pieces of information are unfortunately missing

because residue (i+1) is Hyp^3 and the domination by the extended conformation. It is more surprising that, contrary to the early hypothesis, in which turn-like structure between Ser^6 and Arg^9 were predicted to be important for activities of antagonists[12], residues between Ser^6 and Arg^9 are not involved in a β-turn as evidenced by the absence of both medium-range NOEs and low amide temperature coefficients for the C-terminal residues, although molecular modeling calculations showed that the bulky aliphatic side-chains at positions 7 & 8 have an inherently stabilized β-turn geometry[14,15].

The less active bradykinin antagonist, i.e. peptide **III** showed no evidence to support the turn-like structure between residues 2 and 5. However, it seems there was some rigidity around Gly^4 since its α-protons exhibit a large chemical shift difference as observed in peptides **I** & **II**. Replacing $DCpg^7$ in peptide **II** with Cpg^7 to form peptide **III** makes peptide **III** less structured, which shows the importance of the D-amino acid at position 7. In peptide **III**, the NHs for residues 5 to 8 exhibited NH temperature coefficients that were uniformly similar at about -5.0 ppb/K. Compared to Cpg^8 in peptide **I**, the NH of Cpg^8 in peptide **III** shifted to high field by 0.3 ppm probably due to shielding from two aliphatic rings. For peptides **IV** (very weak antagonist), no structure other than extended can be defined, except NOEs showed some rigidity for residues $DCpg^7$ and Phe^8 (see results).

Based on this study, the correlation between the existence of a turn-like structure from residues 2 to 5 with antagonist activity is found. The turn-like structures are based on the combination of the slow NH temperature coefficients at position 5 (Table 5.2), the NH/NH NOEs between residues 4 & 5 and a medium range Hα(i)-NH(i+2) (Figure 5.4). This is contrary to the early prediction in which the nature of the β-turn between residue 6 and residue 9 was thought to be responsible for changing the activity from an agonist to an antagonist[12]. We

think the NH temperature coefficients are good indicators for turn-like structures, since the stabilizing effect of a hydrogen bond is important for the turn formation in peptides. As can be seen for peptides I to IV, the NH temperature coefficients at position 5 increase, while at position 9 they decrease with a decrease in activity. This indicates that a turn becomes less likely between residues 2 & 5. In a similar study on a bradykinin agonist-antagonist pair (the agonist $DArg^{0}$-Arg^{1}-Pro^{2}-Hyp^{3}-Gly^{4}-Thi^{5}-Ser^{6}-Pro^{7}-Thi^{8}-Arg^{9} and the corresponding antagonist with DPhe in position 7), a type I β-turn was found between residues 6 and 9 in the agonists and some indication for a turn between residues 2 & 5 was found for the antagonist[18]. For this agonist, the NH of Arg^{9} was involved in an H-bond and displayed the lowest temperature coefficients in the peptide. All of these data suggest that the conformational difference between bradykinin agonists and antagonists are due to a turn formed at *different positions* in the peptides.

References

1 Regoli, D. & Barabe, J. C. (1980), *Pharmacol. Rev.* **32**, 1-46.

2 Farmer, S. G. & Burch R. M. (1991), *Bradykinin Antagonists, Basic and Clinical Research*, Burch, R. M. Ed., Marcel Dekker, New York, 1-31.

3 Proud, D., Reynolds, C. J., Lacapra, S., Schotka, A., Lichtenstein, L. M. & Naclerio, R. M. (1988), *Am. Rev. Respir. Dis.* **137**, 613-616.

4 Naclerio, R. M., Proud, D., Lichtenstein, L. M., Sobotka-Kagey, A., Hemdley, J. O., Sorrentino, J. & Gwaltney, J. M. (1988), *J. Infect. Dis.* **157**, 133-142.

5 Stewart, J. M. & Vavrek, R. J. (1991), *Bradykinin Antagonists, Basic and Clinical Research*, Burch, Ronald M. Ed., Marcel Dekker, New York, 51-95.

6 Vavrek, R. J. & Stewart, J. M. (1985), *Peptides* **6**, 161-164.

7 Rose, G. D., Gierasch, L. M. & Smith, J. A. (1985), *Adv. Protein Chem.* **37**, 1-109.

8 Degrado, W. F. (1988), *Adv. Protein Chem.* **39**, 51-124.

9 Vavrek, R. J., Gera, L. & Stewart, J. M. (1991), *Abstract for poster presented at International Kinin Symposium*, Sept. 1991, Munich, Germany.

10 Lembeck, F., Griesbacher, T., Eckhardt, M., Henke, S., Breipohl, G. & Knolle, J. (1991), *Brit. J. Pharmacol.*, **102**, 297-304.

11 Cheronis, J. C., Whalley, E. T., Nguyen, K. T., Eubanks, S. R., Allen, L. G., Duggan, M. J., Loy, S. D., Bonham, K. A. & Blodgett, J. K. (1992), *J. Med. Chem.* **35**, 1563-1572.

12 Kyle, D. J., Hicks, R. P., Blake, P. R., Klimkowski, V. J. (1991), *Bradykinin Antagonists, Basic and Clinical Research*, Burch R. M., Ed., Marcel Dekker, New York, 131-146.

13 Lee, S. C., Russell, A. F. & Laidig, W. D. (1990), *Int. J. Peptide Protein Res.* **35,** 367-377.

14 Kyle, D. J., Martin, J. A., Burch, R. M., Carter, J. P., Lu, S., Meeker, S., Prosser, J. C., Sullivan, J. P., Togo, J., Noronha-Blob, L., Sinsko, J. A., Walters R. F., Whaley, L. W. & Hiner, R. N. (1991), *J. Med. Chem.* **34**, 2649-2653.

15 Kyle, D. J., Martin, J. A., Farmer, S. G. & Burch, R. M. (1991), *J. Med. Chem.* **34**, 1230-1233.

16 Welsh, J. H., Zerbe, O., von Philipsborn, W. & Robinson, J. A. (1992), *FEBS Letters* **297**, 216-220.

17 Liu, X., Scott, P. G., Otter, A. & Kotovych, G. (1992), *Biopolymers* **32**, 119-130.

18 Otter, A., Bigler, P., Stewart, J. M. & Kotovych, G. (1993), *Biopolymers.***32**, 769-780.

19 Hoult, D. I. (1976), *J. Magn. Reson.* **21**, 337-347.

20 Varian, NMR Spectrometer Systems, System Operation Manual, Pub. No. 87-195100-00, Rev. AO689.

21 Varian, NMR Spectrometer Systems, Variable Temperature Unit, Installation and Maintenance Manual, Pub. No. 87-195402-00, Rev. B0790.

22 States, D. J., Haberkorn, R. A. & Ruben, D. J. (1982), *J. Magn. Reson.* **48**, 286-292.

23 Braunschweiler, L. & Ernst, R. R. (1983), *J. Magn. Reson.* **53**, 521-528.

24 Bax, A. & Davis, D. G. (1985), *J. Magn. Reson.* **65**, 355-360.

25 Bothner-By, A. A., Stephens, R. L., Lee, J., Warren, C. D. & Jeanloz, R. W. (1984), *J. Am. Chem. Soc.* **106**, 811-813.

26 Bax, A. & Davis, D. G. (1985), *J. Magn. Reson.* **63**, 207-213.

27 Kessler, H., Griesinger, Ch., Kerssebaum, R., Wagner, K. & Ernst, R. R. (1987), *J. Am. Chem. Soc.* **109**, 607-609.

28 Griesinger, Ch. & Ernst, R. R. (1987), *J. Magn. Reson.* **75**, 261-271.

29 Jeener, J., Meier, R., Bachmann, P. & Ernst, R. R. (1979), *J. Chem. Phys.* **71**, 4546-4553.

30 Kumar, A., Ernst, R. R. & Wüthrich, K. (1980), *Biochem, Biophys. Res. Commun.* **95**, 1-6.

31 Otting, G., Widmer, H., Wagner, G. & Wüthrich, K. (1986), *J. Magn. Reson.* **66**, 187-193.

32 Wüthrich, K. (1986), *NMR in Proteins and Nucleic Acids*, Wiley Interscience, New York.

33 Hruby, V. J. (1974), *Chemistry and Biochemistry of Amino Acids. Peptides and Proteins*, Weinstein, B. Ed., Vol. 3, Marcel Dekker, New York, 1-188.

34 Bax, A., Sklenar, V. & Summers, M. F. (1986), *J. Magn. Reson.* **70**, 327-331.

35 Bax, A. (1989), *Methods in Enzymology* **176**, 151-168.

36 Williamson, M. P. & Neuhaus, D. (1987), *J. Magn. Reson.* **72**, 369-375.

37 Dyson, H. J. & Wright, P. E. (1991), *Annu. Rev. Biophys. Biophys. Chem.* **20**, 519-538.

38 Wilmot, C. M. & Thornton, J. M. (1988), *J. Mol. Biol.* **203**, 221-232.

39 Chou, P. Y. & Fasman, G. D. (1979), *Biophys. J.* **26**, 367-383.

Table 5.1. Peptides Studied.

I	$DArg^0$-Arg^1-Pro^2-Hyp^3-Gly^4-Thi^5-Ser^6-$DCpg^7$-Cpg^8-Arg^9
II	$DArg^0$-Arg^1-Pro^2-Hyp^3-Gly^4-Phe^5-Ser^6-$DCpg^7$-Cpg^8-Arg^9
III	$DArg^0$-Arg^1-Pro^2-Hyp^3-Gly^4-Phe^5-Ser^6-Cpg^7-Cpg^8-Arg^9
IV	Arg^1-Pro^2-Pro^3-Gly^4-Phe^5-Ser^6-$DCpg^7$-Phe^8-Arg^9

Table 5.2. Pharmacological Data and NH Temperature Dependence Coefficients at Positions 5 and 9 of Peptides I-IV.

Peptides	pA_2(Rat uterus)	pA_2(Guinea pig ileum)	NH Coeff. at Position 5 (-ppb/K)	NH Coeff. at Position 9 (-ppb/K)
I	7.9	6.6	2.6	7.7
II	7.9	7.1	2.9	7.7
III	5.1	5.1	4.4	5.3
IV	inactive	I	5.9	4.2

pA_2 = negative logarithm of molar concentration of antagonist to give 50% inhibition.

I = weak inhibitor.

Table 5.3 Proton Chemical Shifts[a], Coupling Constants[b] and NH Temperature Coefficients[c] of the Peptides Studied[d]

Peptide I.

Peptide Residue	NH	α^e	β	γ	other protons	J(Hz)	ΔT (-ppb/K)
DArg[0]	-	3.88	1.85,1.85	1.63,1.63	δ:3.20,3.20	-	-
					εNH:7.78		
Arg[1]	9.00	4.55	1.76,1.76	1.68,1.64	δ:3.12,3.06	-	-
					εNH:7.50		
Pro[2]	-	4.68	2.28,2.28	1.93,1.86	δ:3.84,3.41	-	-
Hyp[3f]	-	4.52	2.26,2.02	4.56	δ:3.79,3.79	-	-
Gly[4]	8.82	4.05,3.82				5.5/6.0	7.4
Thi[5g]	8.30	4.52	3.38,3.26		6.94(3',4')	7.5	2.6
					7.26(5')		
Ser[6]	8.29	4.43	3.81,3.75			7.0	4.1
DCpg[7h]	8.09	4.24	2.26		1.74,1.64,1.54	8.0	7.1
					1.26		
Cpg[8]	8.47	4.14	2.28		1.78,1.70,1.64,1.53	8.0	7.6
					1.35,1.24		
Arg[9]	7.92	4.13	1.85,1.85	1.69,1.60	δ:3.18,3.18	7.5	7.7
					εNH:7.44		

Peptide II.

Peptide Residue	NH	α	β	γ	other protons	J(Hz)	ΔT (-ppb/K)
DArg^0	-	3.90	1.86,1.86	1.62,1.62	δ:3.20,3.20	-	-
					εNH:7.75		
Arg^1	8.99	4.54	1.70,1.76	1.62,1.62	δ:3.08,2.99	-	-
					εNH:7.49		
Pro^2	-	4.70	2.30,2.30	1.97,1.90	δ:3.83,3.44	-	-
Hyp^3	-	4.53	2.25,2.00	4.56	δ:3.80,3.80	-	-
Gly^4	8.76	3.99,3.79				5.8/5.8	7.5
Phe^5	8.26	4.49	3.14,3.03		7.23(2,6,4)	5.9	2.9
					7.28(3,5)		
Ser^6	8.27	4.43	3.80,3.75			7.5	3.4
DCpg^7	8.03	4.24	2.26		1.75,1.65,1.55	8.1	6.3
					1.28,1.25		
Cpg^8	8.47	4.14	2.28		1.79,1.70,1.64,1.55,	7.6	7.5
					1.36,1.26		
Arg^9	7.91	4.14	1.85,1.70	1.60,1.60	δ:3.17,3.17	7.8	7.7
					εNH:7.43		

Peptide III.

Peptide Residue	NH	α	β	γ	other protons	J(Hz)	ΔT (-ppb/K)
DArg^0	-	3.90	1.86,1.86	1.64,1.64	δ:3.21,3.21 εNH:7.62	-	-
Arg^1	8.96	4.57	1.69,1.64	1.78,1.78	δ:3.12,3.12 εNH:7.48	-	-
Pro^2	-	4.71	2.32,2.32	1.97,1.91	δ:3.85,3.48	-	-
Hyp^3	-	4.51	2.27,2.00	4.57	δ:3.80,3.80	-	-
Gly^4	8.73	4.02,3.72				5.5/6.0	7.2
Phe^5	8.19	4.59	3.15,3.00		7.23(2,6,4) 7.28(3,5)	6.0[i]	4.4[i]
Ser^6	8.21	4.40	3.80,3.75			5.5[i]	4.7[i]
Cpg^7	8.18	4.14	2.24		1.77,1.65,1.55 1.33,1.26	7.5[i]	5.9[i]
Cpg^8	8.17	4.10	2.24		1.77,1.65,1.55 1.33,1.26	6.5[i]	4.9[i]
Arg^9	7.86	4.18	1.84,1.84	1.70,1.58	δ:3.17,3.17 εNH:7.46	7.5	5.3

Peptide IV.

Peptide Res[k]	NH	α^e	β	γ	other protons	J(Hz)	ΔT (-ppb/K)
Arg[1]	-	4.28	1.89,1.89	1.69,1.69	δ:3.14,3.14; εNH:7.53	-	-
Pro[2]	-	4.77	2.38,2.38	2.00,1.93	δ:3.73,3.49	-	-
Pro[3]	-	4.33	2.05,2.00	1.85,1.85	δ:3.83,3.66	-	-
Gly[4]	8.46	3.93,3.86				5.0/5.0	5.3
Phe[5]	8.47	4.46	3.14,2.99		ring:7.24	6.0	5.9
Ser[6]	8.29	4.43	3.79,3.71			7.5	3.4
DCpg[7]	7.84	3.98	2.01		1.07,1.55,1.43,1.30 0.72,0.91,1.47[l]	7.0	5.0
Phe[8]	8.79	4.67	3.37,2.77		ring:7.27	8.5	8.9
Arg[9]	7.78	4.18	1.89,1.76	1.61,1.61	δ:3.18,3.18 εNH:7.46	7.5	4.2

a) Chemical shifts are reported in ppm relative to the methyl group of CD_3OH at 3.30 ppm downfield from TMS.

b) Coupling constants (in Hz) refer to $J_{NH-\alpha}$ and were measured from one-dimensional spectra unless indicated otherwise.

c) The temperature coefficients of the amide protons are average values of temperature coefficients measured at seven points between -5 to 25 °C as discussed in the text.

d) The data were recorded at 500 MHz in CD_3OH/H_2O (80/20 v/v%) at 0 °C and at a pH of 5.2 to 5.4. The concentrations for the samples were between 0.9 and 1.4 mmol/L.

e) Pairs of geminal protons x (x=α, β, γ ...) are not assigned stereospecifically. When two distinct lines were observed, the larger chemical shift was arbitrarily assigned to x and the smaller to x'.

f) Hyp refers to *trans*-4-hydroxy-L-proline.

g) Thi refers to β-(2-thienyl)-L-alanine.

h) Cpg refers to α-cyclopentyl-glycine. The protons of the aliphatic ring were not assigned individually, unless otherwise indicated.

i) These values were taken from a 1D spectrum at 25°C. Because of spectral overlap, the error limits may be as large as ± 1 Hz.

j) These values are based on two TOCSY experiments at 0°C and 25°C.

k) Additional weak peaks (~12%) were observed for this peptide.

l) The assignment for the aliphatic ring protons are: 1CH(2.01), $2,5CH_2$(0.72/0.91, 1.07/1.55), $3,4CH_2$(1.47/1.47, 1.43/1.30).

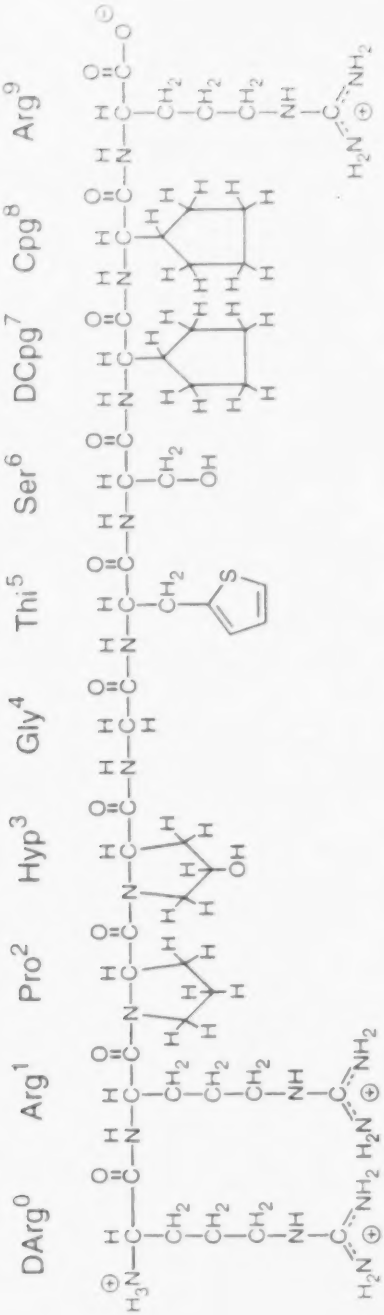


Figure 5.1 The primary structure of peptide I.

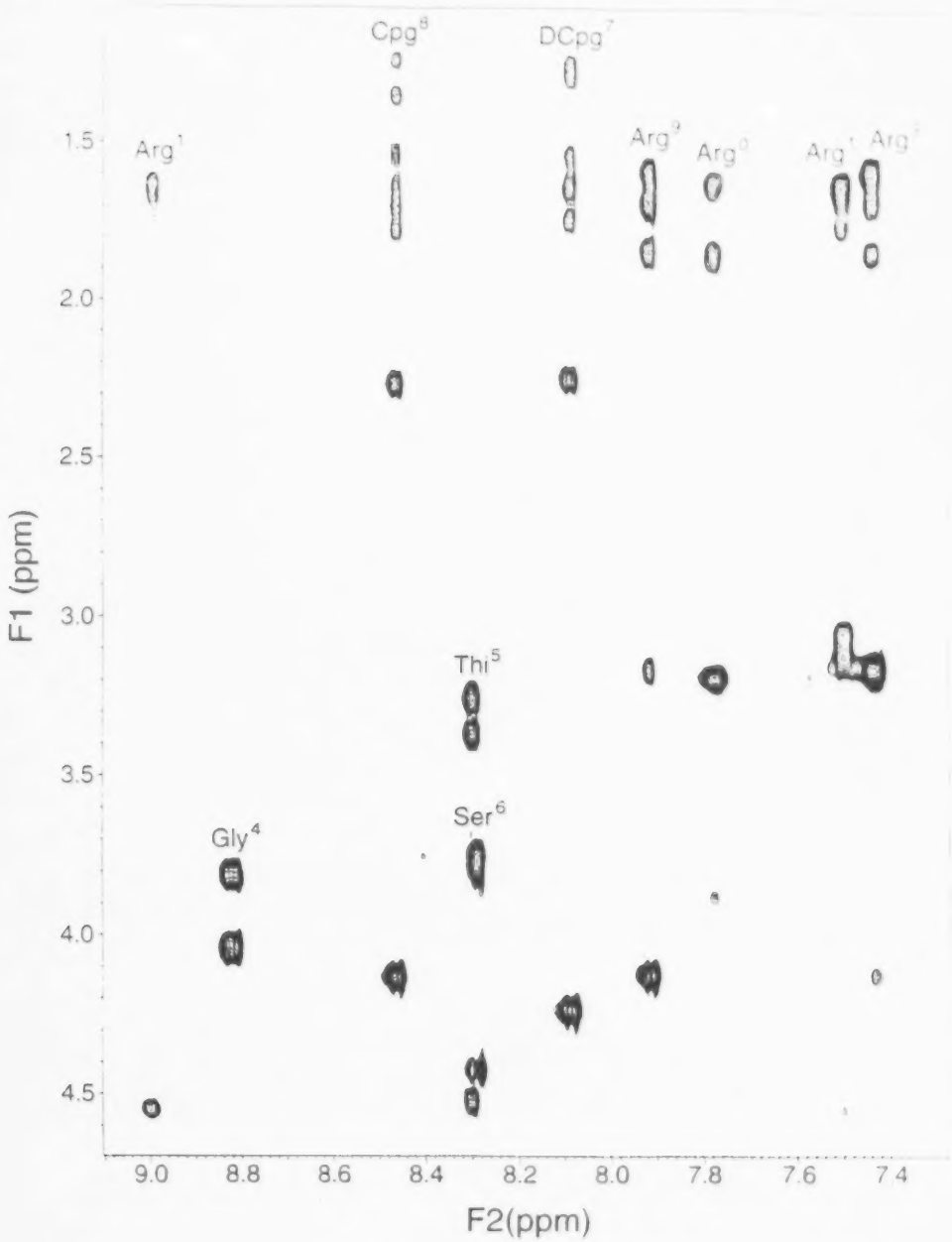


Figure 5.2 The contour plot of a phase-sensitive 500 MHz TOCSY experiment of peptide I in CD_3OH/H_2O(80/20). The concentration was 1.4 mM, pH 5.4 and temperature 0°C. A spin-locking time of 80 ms was used. Experimental details are discussed in the text.

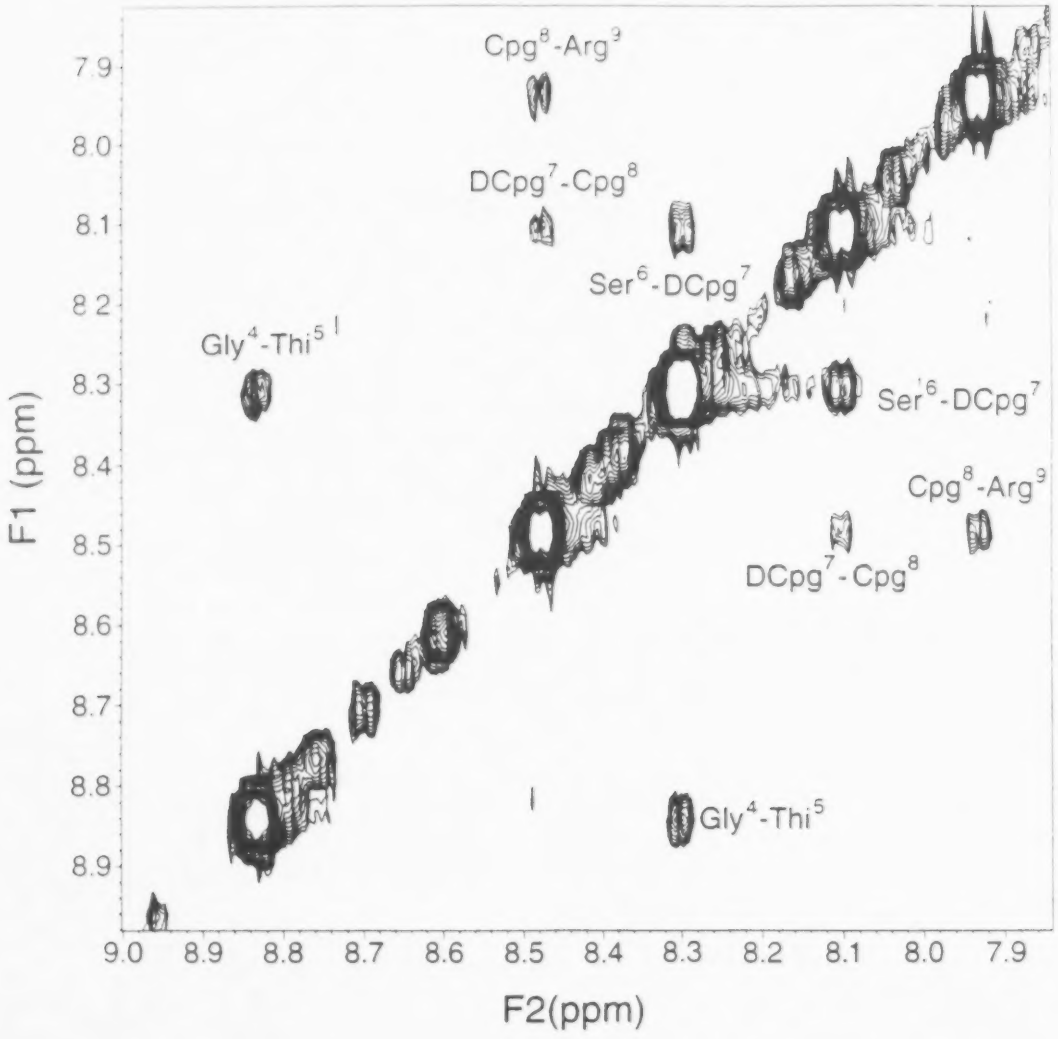


Figure 5.3 The NH/NH region of a 500 MHz phase-sensitive NOESY experiment of peptide I at a mixing time of 300 ms. The experimental conditions are as outlined in Figure 5.2.

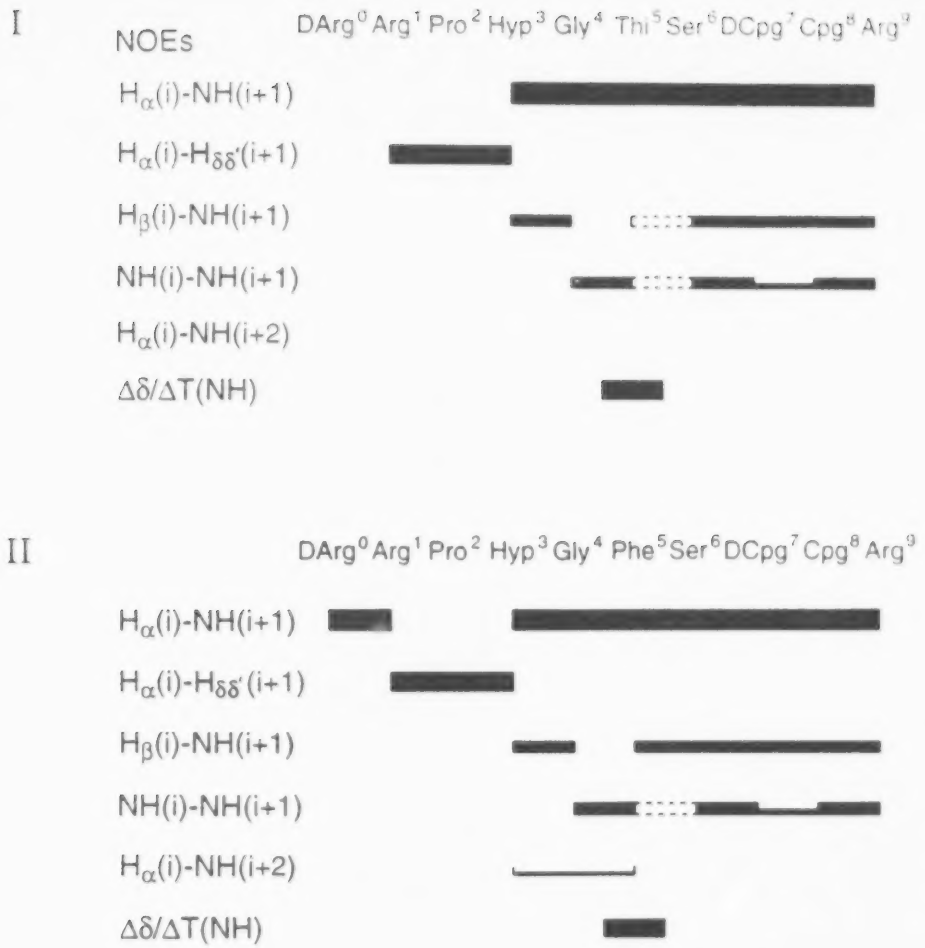


Figure 5.4 The amino acid sequences and the NOE summary of peptides I and II. The observed sequential and medium range NOEs are indicated schematically by solid bars. The height of the bars reflects the NOE intensities based on integration. The temperature coefficients of the amide protons are shown with a solid bar for values between 0 and -3.0 ppb/K. NOEs shown with dashed lines might be present but were not observable due to almost identical chemical shifts of the amide protons or spectral overlap.

CHAPTER 6

The Aggregation Properties of Some Bradykinin Analogs*

Introduction

Bradykinin (BK) is a linear peptide hormone (Arg^1-Pro^2-Pro^3-Gly^4-Phe^5-Ser^6-Pro^7-Phe^8-Arg^9) and has been implicated in a multitude of pathophysiological processes[1,2]. Of particular significance is its role as a potent pain-producing agent. Recently, it has been suggested that bradykinin may be associated with the symptoms of septic shock, asthma and the common cold[3,4]. Because of these physiological functions, a bradykinin receptor antagonist may have significant therapeutic value, and the synthesis of these antagonists has been the subject of intensive research.

Many bradykinin antagonists have been synthesized with the aim of increasing the potency, selectivity and lifetime of these antagonists[5,6]. In almost all of these syntheses, analogs of natural amino acids were used to impose specific conformational restraints[7,8] in order to obtain insight about their bioactive conformation. So far, bulky, beta-branched or cyclic D-aliphatic residues at position 7 combined with bulky or cyclic L-aliphatic residues at position 8 have yielded the best antagonists[9,10].

Conformational studies in solution of a series of bradykinin agonists, antagonists and analogs have been carried out. Our results led to a hypothesis that the difference between an agonist and an antagonist is due to a turn formed

---

* A version of this chapter has been accepted by *J. Biomol. Struct. Dyn.* Liu, X., Stewart, J. M., Gera, L., & Kotovych, G. (1993), *J. Biomol. Struct. Dyn.*, in press.

at different positions in the peptides[11,12]. For antagonists, a turn between residues 2 and 5 is important, while for agonists a type I β-turn between residues 6 and 9 is important. In the course of these studies, an interesting abnormal temperature dependence coefficient was also observed for two of the bradykinin analogs and was attributed to aggregation. In the present study, this phenomenon was investigated by ^{1}H NMR and CD for the peptides DArg-[Hyp^3, Thi^5, DSer^6, DCpg^7, Cpg^8]-BK [I] and DArg-[Hyp^3, DSer^6, DCpg^7, Cpg^8]-BK [II] (Cpg = α-cyclopentylglycine; Hyp = 4-hydroxy-L-proline, Thi = β-(2-thienyl)-L-alanine). The complete sequences for these two molecules are listed in Table 6.1, together with several related analogs, which differ only by one or two residue(s) from these two and have been investigated previously by ^{1}H NMR[12] (X. Liu, J. M. Stewart & G. Kotovych, unpublished data). The solvent for this study is a methanol/water (80/20 v/v) mixture, which is a structure-inducing solvent that has been used in the study of other small peptides[13,14].

Materials and Methods

Materials and Sample Preparation

The peptides were synthesized, purified and assayed in the classical BK assay systems as reported previously[6]. The samples used in all ^{1}H NMR experiments were prepared by dissolving purified solid peptides in 160 μl H_2O. The pH was then adjusted by adding small aliquots of diluted NaOH or HCl, followed by adding 640 μl CD_3OH to make the solvent an 80/20 methanol/water mixture. Before the NMR experiments, the samples were degassed with argon. The sample concentrations were 3.2 mM and 1.1 mM, the pH of the samples was 5.6 and 5.4 for peptides I and II, respectively. Peptide I was diluted by a factor of ten with some small changes in chemical shifts.

NMR Spectroscopy

All experiments were carried out on the Varian UNITY 500 NMR spectrometer using a proton-selective 5 mm probe with a 90° proton pulse length of 8.8 to 10.0 μs at normal power levels. During the experiments the temperature was always controlled to ±0.1°C and set to 0°C except for the temperature studies. For FID acquisitions and data processing, VNMR 3.1 and 3.2 software were used on a SUN 4/330 and a SUN 4/470 workstation, respectively. The proton chemical shifts were measured relative to the undeuterated fraction of the methyl group of CD_3OH (Merck, Sharp & Dome, MD-67, 99.9% D) at 3.30 ppm. All 2D spectra were acquired non-spinning, whereas for 1D experiments the spinner was turned on.

In all NMR experiments, the water resonance from the solvent was suppressed during the relaxation delay by presaturation. The decoupler offset was set equal to the transmitter offset to prevent phase distortions around the water peak due to imperfect cancellation of the dispersive component of residual water magnetization[15]. Before each experiment, the delays before and after the reading pulse (alfa delay) were carefully adjusted to give a flat and undistorted baseline[16].

One-dimensional proton spectra were recorded with a sweep width of 5000Hz, a relaxation time of 2s and 16K (zero-filled to 32K) data points. Usually a total of 32 scans was accumulated, and the resulting FID's were resolution enhanced by means of a Gaussian multiplication with a line broadening factor of typically -3 Hz.

The temperature dependence of the amide protons was calculated by measuring the chemical shifts at -5, 0, 5, 10, 15, 20 and 25 °C for peptide **I** and at 10 temperatures between -5 and 40 °C for peptide **II**, then analyzed by means of a linear regression. The controller temperatures were calibrated using a

methanol sample provided by Varian[17] and were found to be in the range of +/- 1.0 °C of the predicted temperature.

Two-dimensional experiments were measured with 2048 (zero filled to 4096) or 4096 data points along t_2 and 256 t_1 increments. Spectral widths of 5000 Hz in both dimensions and a relaxation delay of 1.5 to 2.0 s were employed. All two-dimensional experiments were acquired in the phase-sensitive mode by using the hypercomplex technique known as the States-Haberkorn-Ruben method[18]. The TOCSY[19] experiments were carried out by using the basic pulse sequence proposed by Bax[20]. A 2ms trim pulse preceded the 70-80ms MLEV-17 spin-lock (field strength 7.0 kHz). 16 to 32 scans were accumulated per experiment. The NOESY experiments[21,22] were carried out with the 90°-t1-90°-mix-90°-FID pulse sequence, 200 or 300 ms mixing time and 32 scans per t1 increment. Water suppression was achieved by decoupler saturation at all times except during the acquisition period. A 10% random delay was incorporated in the mixing delay to suppress zero quantum coherences. For TOCSY, and NOESY experiments, the compensation for first point distortion caused by analog filters[23] was achieved by reducing the first point of each FID and each t1-interferogram by empirically determined values[16]. In all cases, a $\pi/2$ shifted, squared sine bell weighting function was used in both dimensions. For all the 2D spectra, baseline corrections were not necessary.

For the deuterium exchange experiments, peptides I & II were freeze-dried and the deuterium solvent was added into the NMR tube just before the tube was placed into the magnet. Normal one-D NMR spectra were recorded regularly until all the NH resonances disappeared.

Circular Dichroism

The CD spectra (Figure 6.3) were recorded at 10±0.05 °C for peptide **I** and β-(2-thienyl)-L-alanine (27°C for β-(2-thienyl)-L-alanine methyl-amide) on a Cary model 60 spectropolarimeter with a model 6001 circular dichroism attachment with thermostated cell holder. The instrument had been modified to eliminate possible artifactual signals on passing through intense absorption bands. Slits were programmed for 1.5 nm band width at each wavelength. The instrument was calibrated with dextro-10-camphor-sulfonic acid using the procedure suggested by Chen and Yang[24]. The path length, 0.02 cm to 2 cm depending upon the concentration and the wavelength range, was dictated by the absorbance of the solution. Solutions were prepared by dissolving weighed quantities of peptide **I** in 80/20 methanol/water, and their concentrations calculated by using the peptide percentage and the molecular weight of the peptide. Mean residue ellipticities, $[\theta]$ mrw (deg cm^2 $dmol^{-1}$), were calculated in the usual fashion using the mean residue weight, 125, reckoned from the peptide's molecular weight. Each spectrum is the average of two rescans. UV absorption spectra were recorded at 27°C on a Beckman model 25 spectrophotometer.

Results and Discussion

Chemical Shift Assignments and Coupling Constants

The phase-sensitive TOCSY spectra of the two peptides in methanol/water (80/20) alone allowed the identification of all unique amino acids in a sequence[23]. Then, a sequential analysis was performed using the NOESY[25] to achieve the complete assignment. These assignments rely on H_α-NH(i, i+1), NH-NH(i, i+1) and for Pro residues, H_α-$H_{\delta,\delta'}$(i, i+1) connectivities. The NH resonance of the first N-terminal residue is always absent due to fast

exchange with the solvent since the peptides are not protected. In addition, NH resonances of Arg[1] for peptides are usually very weak and can only be seen at low temperature. However, the spin systems for these Arg resonances can be easily identified through their NH resonance in the side chain. It was difficult to assign specifically each of the aliphatic ring protons in the Cpg residues because of spectral overlap. The cis isomers due to the X-Pro or X-Hyp peptide bond do exist, but in very small amounts (less than 10%). This observation is consistent with the suggestion based on NMR experiments run in a lipophilic environment that all amide bonds remain trans in both agonists and antagonists[26].

The assigned resonances for two peptides are summarized in Table 6.2. In order to show the effect of a DSer[6] residue on the peptide NH chemical shifts, the chemical shifts for the NHs and the α-protons for the L-Ser[6] analogs are also listed in the table[12].

Peptides **I** and **II** exhibit many similarities in terms of chemical shifts, which is understandable since the peptides only differ by one residue (Thi[5] in **I** is replaced by Phe[5] in **II**), as do the two L-Ser[6] analogs. From peptide **I** to peptide **II**, very little change in the chemical shifts of the amide protons was observed. However, it is interesting to notice that the chemical shift for some NH protons changes when DSer[6] in peptides **I** and **II** is replaced by L-Ser[6]. The chemical shifts for the NH protons of Cpg[8] shift to low field (7.86 to 8.47 for **I**, 7.77 to 8.47 for **II**). The chemical shifts of the other NH resonances for the neighboring residues were also changed, but to a lesser extent. Although changes in chemical shifts are not necessarily a good probe for conformational changes in peptides, they are a sensitive indicator for the changes in the chemical environment which these NH protons experience.

All vicinal coupling constants of the type $^3J_{\alpha NH}$ could be measured in the one-dimensional spectra, at different temperatures as required to minimize

spectral overlap. None of the coupling constants changed by more than 0.3 Hz over the temperature range -5 to 25°C, an indication that the conformation does not change significantly. All the observed $^3J_{\alpha NH}$ values are close to the time-averaged value of 6.3 Hz for free rotation about the N-C_α bond.

Temperature Dependence of NH-chemical Shifts

An amide proton involved in a stable intramolecular hydrogen bond, or one inaccessible to solvent for steric reasons, typically shows a reduced temperature dependence coefficient in the range 0 to -3 ppb/K[27]. That means that with an increase in the temperature, the NH proton shifts should move toward high field since the signal for the water residue moves to high field. In this study, the temperature range for measuring the temperature dependence coefficients of the amide protons was between -5 and 25 °C for peptide **I** and between -5 and 40°C for peptide **II**. The results showed that the chemical shifts for all the amide protons except the amide protons of Cpg^8 & Arg^9 varied linearly with temperature. More surprisingly, the amide proton of Cpg^8 shifted to lower field initially with an increase in temperature, in the opposite direction to those of the shifts of the other NH protons. The temperature coefficient changed from +1.4 between -5 and 0 °C to -0.8 between 35 and 40 °C (Figure 6.1). These results suggest that at a lower temperature, the NH of Cpg^8 is shielded from the solvent by the two bulky aliphatic rings and may also exchange with other NH protons at lower field. In the case of Arg^9, the NH temperature coefficient displayed linearity initially at lower temperatures, but showed non-linearity at higher temperatures (above 25°C) (Figure 6.1). The temperature coefficients for the amide protons of these two peptides are listed in Table 6.2. The abnormal temperature dependence coefficient for residue 8 and the non-linearity for residue 9 were not observed on peptides **III** to **IX**[12] (X. Liu, J. M. Stewart and G.

Kotovych, unpublished results), for which the temperature dependence coefficients were obtained between -5 and 25 °C. Based on these observations (Table 6.2), the abnormal temperature dependence behavior of peptides **I** & **II** is only present for the sequence of DSer^6-DCpg^7-Cpg^8-Arg^9, i.e., D^6-D^7-L^8-L^9 is the required sequence of configurations of the C-terminal amino acids.

Isotope Exchange

Figure 6.2 shows the behavior of the NH protons of Arg^9, DCpg^7 and Cpg^8 for peptide **I** in the CD_3OD/D_2O (80/20) solvent. The first spectrum in the figure was recorded in CD_3OH/H_2O (80/20) solvent, which displays a small difference in chemical shift compared to the rest of the spectra due to the solvent isotope effect[28]. With these three exceptions, the remainder of the NH proton resonances disappeared before the first spectrum was recorded 2 minutes after the deuterated solvent was added. As can be seen, the isotope exchange rates for the three NHs are very slow, considering the size of the peptide. The exchange rate for the NH of the Arg^9 is slower, and the NH of DCpg^7 is faster than the NH of Cpg^8, which displayed the abnormal temperature dependence coefficient. It seems that the NHs of the three residues are trapped and not exposed to the bulk solvent in the aggregate. These aggregation properties are also supported by the CD data where aggregation is already observed for a 0.150 mM solution (Figure 6.3). Similar behavior was observed for peptide **II** in the isotope exchange experiment, except the NHs of DCpg^7, Cpg^8, and Arg^9 exchanged more rapidly because of a lower sample concentration. Isotope exchange experiments were also carried out on peptides **IV** & **VII**, and slow exchange rates for the NH protons in these two molecules were not observed.

The salient feature of the CD spectrum of peptide **I** (Figure 6.3) is the broad positive band extending from 260 nm down to 213 nm with an extremum

at 225 nm. This band would not appear to be conformationally related seeing that β- and γ-turns generally exhibit negative CD bands in this spectral region even when the peptide contains D-aminoacids[29,30]. Although in a few cases the band is positive, the extrema are located much further into the blue[31,32]. On the basis of the spectral properties of the model compound β-(2-thienyl)-L-alanine methylamide (inset to Figure 6.3), we tentatively assign the band to the thienyl moiety of the Thi aminoacid residue in the peptide: the model compound exhibits a broad UV absorption band with λ_{max}=234 nm characteristic of thiophene[33] and , as anticipated, the band is optically active. In particular, note the broad system of strongly overlapping CD bands extending from 260 nm to 218 nm. It is interesting that the positive band in the peptide spectrum is insensitive to concentration, suggesting that the same chiral environment of the thienyl moiety is maintained over the designated concentration range.

In summary, our experimental results, i.e., abnormal temperature dependence of the NH protons of Cpg^8, the slow exchange rate of the NH protons of Arg^9, $DCpg^7$ and Cpg^8, and the CD spectrum, suggest that aggregation exists around the Cpg^8 residue in peptides **I** & **II**. This aggregation was observed only for peptides with the sequence $DSer^6$-$DCpg^7$-Cpg^8-Arg^9, as only these two peptides displayed the abnormal temperature dependence (Table 6.1). Our model for this aggregation is that two molecules interact in a head-to-tail fashion, with an overlap of residues 6 to 9. In this way, the hydrophilic side-chains of Ser^6 and Arg^9 point approximately away from the hydrophobic core, while the highly hydrophobic side-chains of the two Cpg residues point inwards, forming the hydrophobic core. It is in this core that the NH of Cpg^8 is trapped between two bulky aliphatic rings of residues $DCpg^7$ and Cpg^8, which causes the inaccessibility of this NH to the solvent. The driving force for this aggregation is likely the hydrophobic interaction. Some inter-residue NOEs between the NHs

and side-chains were observed for residues 6 to 9, but spectral overlap in the aliphatic region, together with aggregation, prevent us from drawing any conformational conclusion at the concentrations used for the NMR experiments.

Aggregation is of concern in all spectroscopic studies of peptides. The conclusions that can be drawn about peptide conformational preferences depend critically upon whether a structure is intrinsic to the monomer or is induced by aggregation[34]. Usually, measurement of the concentration dependence of the NMR or CD spectra (over the widest possible concentration range) is sufficient to identify peptides that have a tendency to associate. In the case of the two peptides under study, the concentration of the samples was low, but it was the abnormal temperature dependence coefficients that warned us of aggregation, which was confirmed further by the exchange and the CD measurements.

References

1 Regoli, D. & Barabe, J. C. (1980), *Pharmacol. Rev.* **32**, 1-46.

2 Farmer, S. G. & Burch R. M. (1991), *Bradykinin Antagonists, Basic and Clinical Research*, Burch, R. M. Ed., Marcel Dekker, New York, 1-31.

3 Proud, D., Reynolds, C. J., Lacapra, S., Schotka, A., Lichtenstein, L. M. & Naclerio, R. M. (1988), *Am. Rev. Respir. Dis.* **137**, 613-616.

4 Naclerio, R. M., Proud, D., Lichtenstein, L. M., Sobotka-Kagey, A., Hemdley, J. O., Sorrentino, J. & Gwaltney, J. M. (1988), *J. Infect. Dis.* **157**, 133-142.

5 Stewart, J. M. & Vavrek, R. J. (1991), *Bradykinin Antagonists, Basic and Clinical Research*, Burch, Ronald M. Ed., Marcel Dekker, New York, 51-95.

6 Vavrek, R. J. & Stewart, J. M. (1985), *Peptides* **6**, 161-164.

7 Rose, G. D., Gierasch, L. M. & Smith, J. A. (1985), *Adv. Protein Chem.* **37**, 1-109.

8 Degrado, W. F. (1988), *Adv. Protein Chem.* **39**, 51-124.

9 Vavrek, R. J., Gera, L. & Stewart, J. M. (1991), *Abstract for poster presented at International Kinin Symposium*, Sept. 1991, Munich, Germany.

10 Lembeck, F., Griesbacher, T., Eckhardt, M., Henke, S., Breipohl, G. & Knolle, J. (1991), *Brit. J. Pharmacol.* **102**, 297-304.

11 Otter, A., Bigler, P., Stewart, J. M. & Kotovych, G. (1993), *Biopolymers* **32**, 769-780.

12 Liu, X., Stewart, J. M., Gere, L., & Kotovych; G. (1993), *Biopolymers*, in press.

13 Liu, X., Scott, P. G., Otter, A. & Kotovych, G. (1992), *Biopolymers* **32**, 119-130.

14 Liu, X., Scott, P. G., Otter, A. & Kotovych, G. (1990), *J. Biomol. Struct. Dyn.* **8**, 63-80.

15 Hoult, D. I. (1976), *J. Magn. Reson.* **21**, 337-347.

16 Varian, NMR Spectrometer Systems, System Operation Manual, Pub. No. 87-195100-00, Rev. AO689.

17 Varian, NMR Spectrometer Systems, Variable Temperature Unit, Installation and Maintenance Manual, Pub. No. 87-195402-00, Rev. B0790.

18 States, D. J., Haberkorn, R. A. & Ruben, D. J. (1982), *J. Magn. Reson.* **48**, 286-292.

19 Braunschweiler, L. & Ernst, R. R. (1983), *J. Magn. Reson.* **53**, 521-528.

20 Bax, A. & Davis, D. G. (1985), *J. Magn. Reson.* **65**, 355-360.

21 Jeener, J., Meier, R., Bachmann, P. & Ernst, R. R. (1979), *J. Chem. Phys.* **71**, 4546-4553.

22 Kumar, A., Ernst, R. R. & Wüthrich, K. (1980), *Biochem, Biophys. Res. Commun.* **95**, 1-6.

23 Otting, G., Widmer, H., Wagner, G. & Wüthrich, K. (1986), *J. Magn. Reson.* **66**, 187-193.

24 Chen, G.C. & Yang, J.T. (1977), *Anal. Letters* **10**, 1195-1207.

25 Wüthrich, K. (1986), *NMR in Proteins and Nucleic Acids*, Wiley Interscience, New York.

26 Kyle, D. J., Martin, J. A., Farmer, S. G. & Burch, R. M. (1991), *J. Med. Chem.* **34**, 1230-1233.

27 Hruby, V. J. (1974), *Chemistry and Biochemistry of Amino Acids. Peptides and Proteins*, Weinstein, B. Ed., Vol. 3, Marcel Dekker, New York, 1-188.

28 Otter, A., Liu, X. & Kotovych, G. (1990), *J. Magn. Reson.* **86**, 675-662.

29 Yang J. T., Wu, C-S C. & Martinez, H. M. (1986), *Methods in Enzymology* **130**, 208-269.

30 Madison, V., Atreyi, M., Deber, C. M. & Blout, E. R. (1974), *J. Am. Chem. Soc.* **96**, 6725-6734.

31 Gierasch, L. M., Deber, C. M., Madison, V., Niu, C-H & Blout, E. R. (1981), *Biochemistry* **20**, 4730-4738.

32 Ananthanarayanan, V. S. & Shyamasundar N. (1981), *Biochem. Biophys. Res. Commun.* **102**, 295-301.

33 Jaffé, H. H. & Orchin, M. (1962), *Theory and Applications of Ultraviolet Spectroscopy*, John Wiley & Sons, New York, 347-348.

34 Dyson, H. J. & Wright, P. E. (1991), *Annu. Rev. Biophys. Biophys. Chem.* **20**, 519-538.

Table 6.1 Peptides Studied.

I	$DArg^{0}$-Arg^{1}-Pro^{2}-Hyp^{3}-Gly^{4}-Thi^{5}-$DSer^{6}$-$DCpg^{7}$-Cpg^{8}-Arg^{9}
II	$DArg^{0}$-Arg^{1}-Pro^{2}-Hyp^{3}-Gly^{4}-Phe^{5}-$DSer^{6}$-$DCpg^{7}$-Cpg^{8}-Arg^{9}
III	$DArg^{0}$-Arg^{1}-Pro^{2}-Hyp^{3}-Gly^{4}-Thi^{5}-Ser^{6}-$DCpg^{7}$-Cpg^{8}-Arg^{9}
IV	$DArg^{0}$-Arg^{1}-Pro^{2}-Hyp^{3}-Gly^{4}-Phe^{5}-Ser^{6}-$DCpg^{7}$-Cpg^{8}-Arg^{9}
V	$DArg^{0}$-Arg^{1}-Pro^{2}-Hyp^{3}-Gly^{4}-Phe^{5}-Ser^{6}-Cpg^{7}-Cpg^{8}-Arg^{9}
VI	$DArg^{0}$-Arg^{1}-Pro^{2}-Hyp^{3}-Gly^{4}-Thi^{5}-$DSer^{6}$-$DCpg^{7}$-$DCpg^{8}$-Arg^{9}
VII	$DArg^{0}$-Arg^{1}-Pro^{2}-Hyp^{3}-Gly^{4}-Thi^{5}-Ser^{6}-$DCpg^{7}$-$DCpg^{8}$-Arg^{9}
VIII	Arg^{1}-Pro^{2}-Pro^{3}-Gly^{4}-Phe^{5}-$DSer^{6}$-$DCpg^{7}$-Phe^{8}-Arg^{9}
IX	Arg^{1}-Pro^{2}-Pro^{3}-Gly^{4}-Phe^{5}-Ser^{6}-$DCpg^{7}$-Phe^{8}-Arg^{9}

Table 6.2 Proton Chemical Shifts[a], Coupling Constants[b] and NH Temperature Coefficients[c] of the Peptides Studied[d]

Peptide I.

Peptide Residue	NH	α[e]	β	γ	other protons	J(Hz)	ΔT (-ppb/K)
DArg^0	-	3.92	1.86,1.86	1.62,1.62	δ:3.21,3.21	-	-
					εNH:7.64		
Arg^1	8.96	4.59	1.84,1.84	1.70,1.61	δ:3.14,3.14	-	-
	(9.00)[f]	(4.55)			εNH:7.62		
Pro^2	-	4.70	2.30,2.30	2.00,1.89	δ:3.87,3.50	-	-
Hyp^3[g]	-	4.62	2.22,2.00	4.52	δ:3.81,3.81	-	-
Gly^4	8.91	3.94,3.94				6.0/5.5	7.4
	(8.82)	(4.05,3.82)					
Thi^5[h]	8.40	4.46	3.30,3.30			5.0	4.8
	(8.30)	(4.52)					
DSer^6	8.56	4.24	3.83,3.52			7.0	6.8
	(8.29)	(4.43)	(3.81,3.75)				
DCpg^7[i]	7.99	4.08	2.28		1.82,1.62,1.53,1.35,	7.5	3.1
	(8.09)	(4.24)			1.26		
Cpg^8	7.86	4.16	2.20		1.77,1.68,1.61,1.53,	8.0	not linear[j]
	(8.47)	(4.14)			1.32, 1.20		
Arg^9	8.03	4.13	1.84,1.84	1.76,1.61	δ:3.17,3.17	7.0	7.4[j]
	(7.92)	(4.13)			εNH:7.46		

Peptide II

Peptide Residue	NH	α	β	γ	other protons	J(Hz)	ΔT (-ppb/K)
DArg^0	-	3.90	1.85,1.85	1.60,1.60	δ:3.19,3.19	-	-
					εNH:7.64		
Arg^1	8.95	4.59	1.75,1.75	1.57,1.57	δ:3.10,3.10	-	-
	(8.99)	(4.54)			εNH:7.68		
Pro^2	-	4.71	2.30,2.30	2.00,1.90	δ:3.84,3.48	-	-
Hyp^3	-	4.62	2.18,1.96	4.53	δ:3.80,3.80	-	-
Gly^4	8.94	3.96,3.91				5.5/6.0	7.9
	(8.76)	(3.99,3.79)					
Phe^5	8.44	4.40	3.08,3.03		7.23(2,6)	5.0	5.8
	(8.26)	(4.49)			7.32(3,5), 7.26(4)		
DSer^6	8.45	4.16	3.74,3.36			8.0	6.2
	(8.27)	(4.43)					
DCpg^7	7.99	4.04	2.27		1.80,1.24(2)	7.5	2.9
	(8.08)	(4.24)			1.61,1.35(5)		
					1.52		
Cpg^8	7.77	4.15	2.18		1.77,1.18(2)	8.0	not linear
	(8.47)	(4.14)			1.66,1.32(5)		
					1.53		
Arg^9	8.05	4.12	1.82,1.82	1.70,1.61	δ:3.17,3.17	7.5	7.4
	(7.91)	(4.14)			εNH:7.46		

a) Chemical shifts are reported in ppm relative to the methyl group of CD_3OH at 3.30 ppm downfield from TMS.

b) Coupling constants (in Hz) refer to $J_{NH\text{-}\alpha}$ and were measured from one-dimensional spectra.

c) The temperature coefficients of the amide protons are average values of temperature coefficients measured at seven points between -5 to 25 °C as discussed in the text.

d) The data were recorded at 500 MHz in CD_3OH/H_2O (80/20 v/v%) at 0 °C. The concentrations were 3.2 and 1.1 mmol/L, the pH 5.4 and 5.6 for peptides **I** and **II**, respectively.

e) Pairs of geminal protons x (x=α, β, γ ...) are not assigned stereospecifically. When two distinct lines were observed, the larger chemical shift was arbitrarily assigned to x and the smaller to x'.

f) The numbers in brackets are the chemical shifts for the corresponding L-Ser peptide.

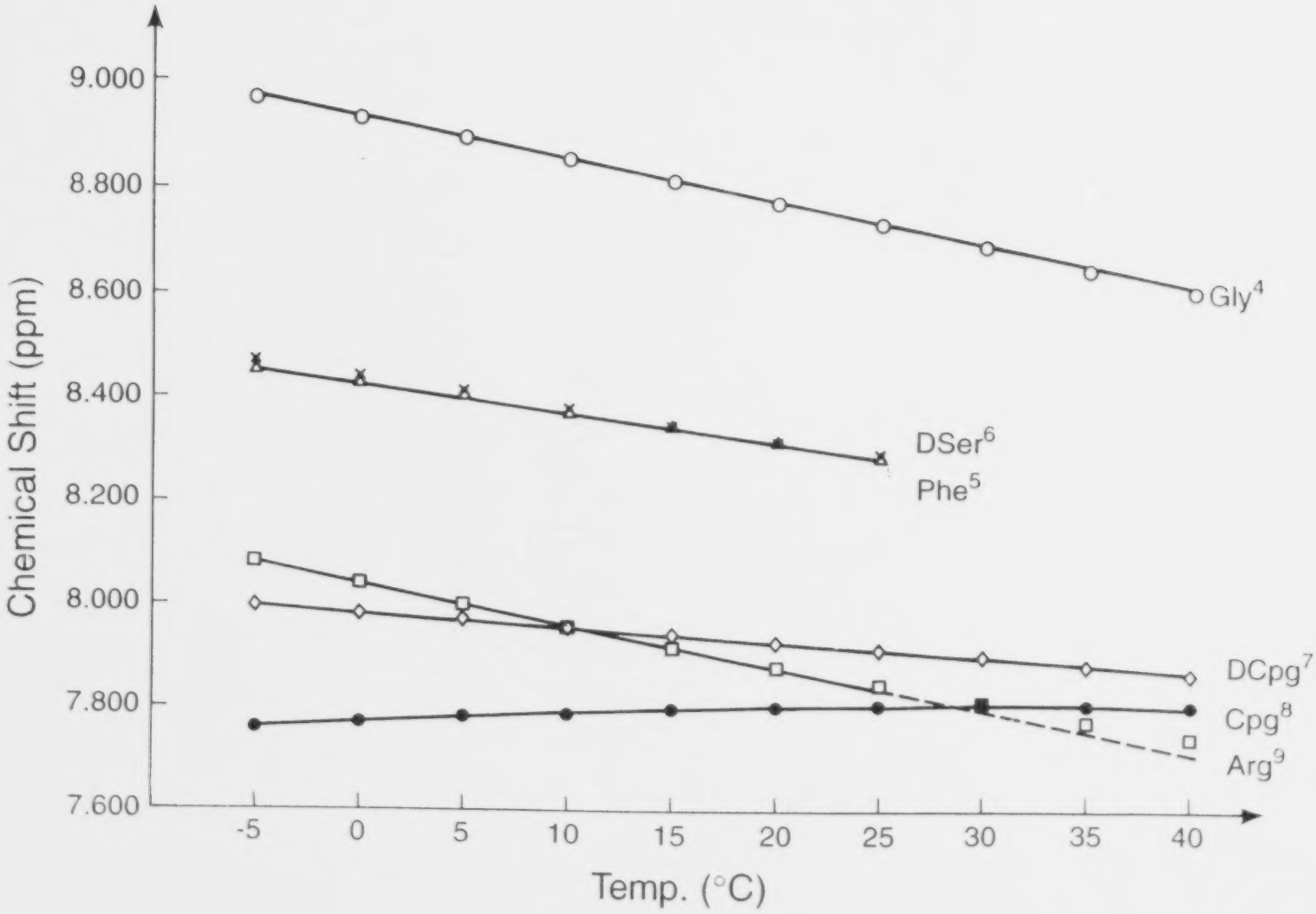
g) Hyp refers to 4-hydroxy-L-proline.

h) Thi refers to β-(2-thienyl)-L-alanine.

i) Cpg refers to α-cyclopentyl-glycine. The protons of the aliphatic ring were not assigned individually because of spectral overlap.

j) The temperature dependence coefficients showed non-linearity. See Figure 1 and text for detail.

Figure 6.1 Plot of the chemical shifts of the amide protons of peptide II in CD_3OH/H_2O(80/20) vs temperature. The concentration was 1.1 mM and pH 5.4. Identical acquisition and processing parameters were used for all spectra. The NH of Cpg^8 initially shifts to low field and then shifts to high field when the temperature is changed from -5 to 40°C, while the other amide protons shift to high field. The NH shift of Arg^9 is non-linear above 20°C. Details are discussed in the text.


9.000
8.800
8.600
8.400
8.200
8.000
7.800
7.600
Chemical Shift (ppm)
-5
0
5
10
15
20
25
30
35
40
Temp. (°C)
Gly4
DSer6
Phe5
DCpg7
Cpg8
Arg9

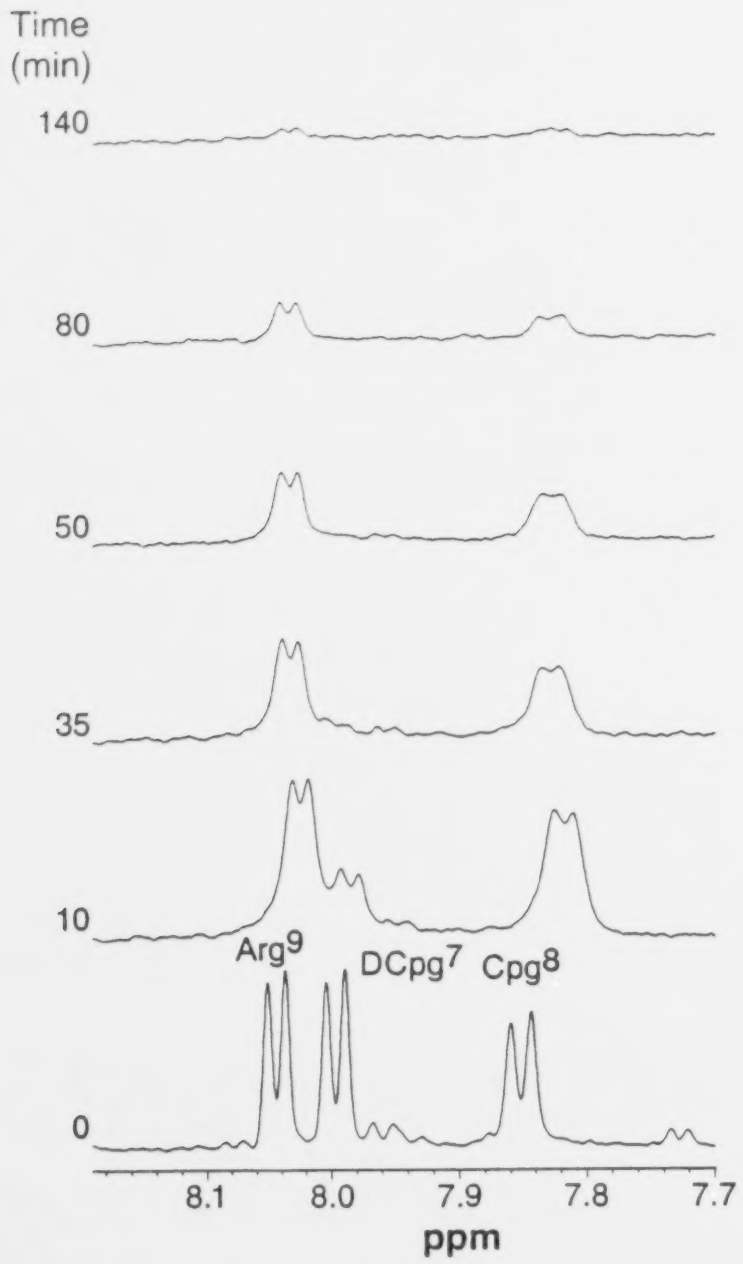
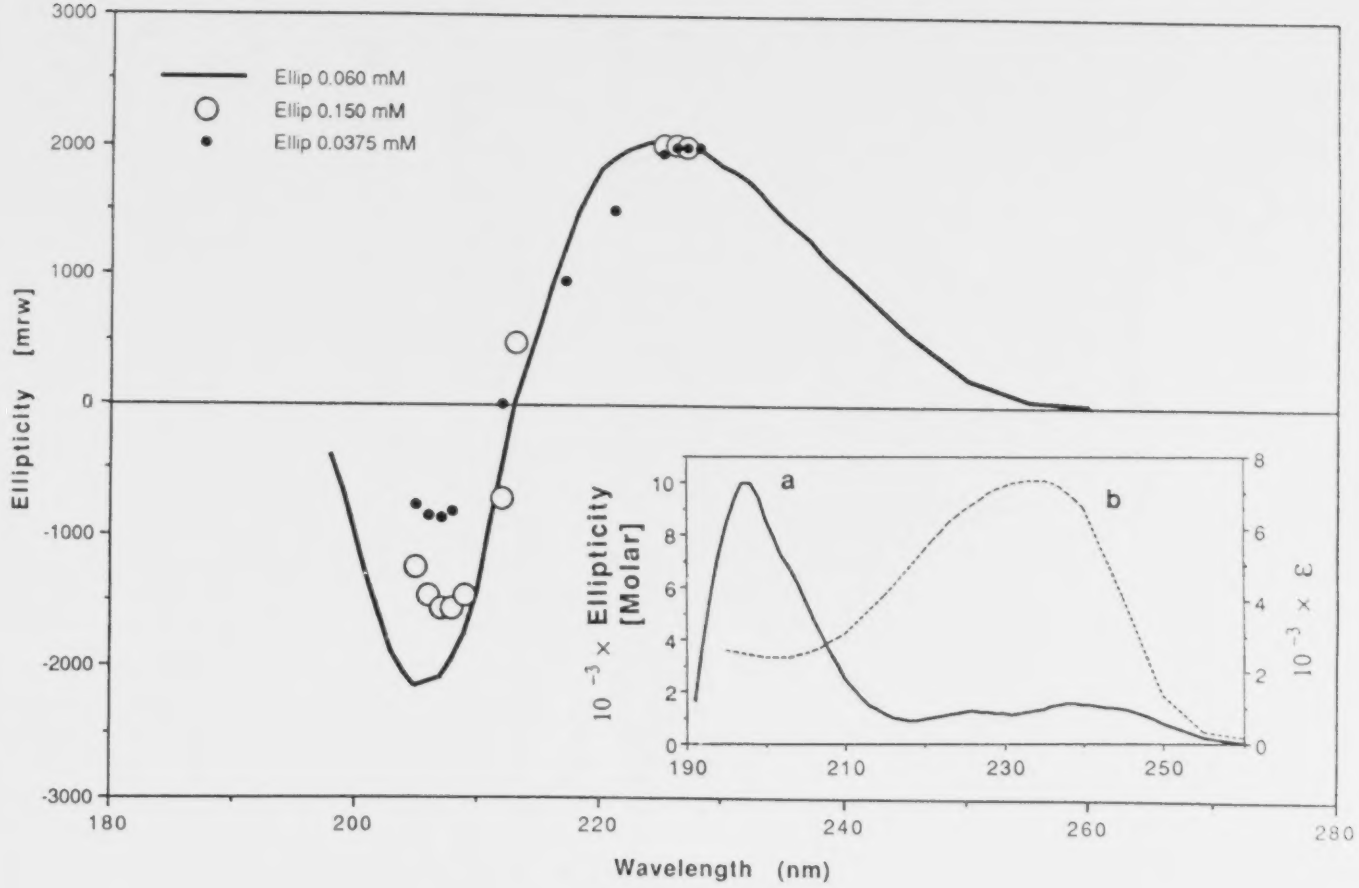


Figure 6.2 Plot of part of the NH region of peptide **I** in $CD_3OD/D_2O(80/20)$ vs time. The concentration was 3.2 mM and pH 5.6. The temperature was 0°C. The spectrum at time 0 was recorded in $CD_3OH/H_2O(80/20)$. The other individual spectra were recorded at the indicated times after sample preparation.

Figure 6.3 Circular dichroism spectrum of peptide I in different concentrations in methanol/water (80/20). The pH of the sample was about 4. Inset, spectral properties of β-(2-thienyl)-L-alanine methylamide: a, CD spectrum; b, UV absorption spectrum. Similar spectra for β-(2-thienyl)-L-alanine.


Ellip 0.060 mM
Ellip 0.150 mM
Ellip 0.0375 mM
Ellipticity [mrw]
Wavelength (nm)
3000
2000
1000
0
-1000
-2000
-3000
180
200
220
240
260
280
a
b
10 -3 × Ellipticity [Molar]
10 -3 × ε
190
210
230
250

CHAPTER 7

General Discussion and Conclusions

The conformational analysis of several small linear peptides, using high-field NMR spectroscopy and CD spectra, has been presented in this thesis. These conformational analyses are useful for understanding the relationship between the conformation and the biological activities, as well as the stabilizing factors on the local conformation and the mechanism of the early steps of protein folding.

In recent years, both NMR and CD techniques for protein conformational studies have been advancing very rapidly[1,2]. This advance is the result of hardware development, increasing computer power, biological techniques to label proteins with ^{15}N and ^{13}C isotopes, and the nature of protein molecules themselves, which display unique and densely packed conformations and yield a great deal of information through NOEs, coupling constants and NH exchange rates. Now, proteins with 200-300 residues can be readily studied by NMR.

In principle, new techniques that are applied to proteins can also be applied to small peptides. This does not mean that the conformational studies for small peptides are easier. Small peptides have far fewer structural restraints and every atom in the molecule is usually exposed to the solvent, which means that less information is available from NMR experiments. Most importantly, peptides are flexible and adopt multiple conformations, the consequence of which is that the NMR parameters are time-averaged over each individual conformation. Because of these properties of peptides, the general procedure for proteins is limited in its application to peptides[3].

For bioactive peptides, the choice of solvent is an important issue for two reasons. First of all, a proper solvent should mimic the environment experienced by the ligand in the presence of the receptor. It has been suggested that receptor binding sites are hydrophobic. In addition, many peptide hormones bind to membrane-bound receptors. This would suggest that solvents with low dielectric constants should be the most realistic. However, in apolar solvents, structures with too many internal hydrogen bonds could be stabilized. In terms of simulation of an *in vivo* environment, a micellar solution could be a better choice of solvent[15].

Secondly, the proper solvent can be used to enhance secondary structures. At present, almost all studies on proteins are in aqueous solutions and some small linear peptides have been found to adopt folded conformations in aqueous solutions[4]. However, even in the cases that the peptides were found to adopt some secondary structures, the folded structures are less stable and less ideal, and the percentage of the folded structures is usually low[4,5]. On the other hand, the structural stability of marginally stable small peptides and the physical interactions accounting for the stability are of current interest in order to understand the biological actions of peptides as well as the mechanism of early steps of protein folding. Under these circumstances, some less polar solvents are being used as stabilizing co-solvents to enhance the secondary structures. The most popular one is trifluoroethanol (TFE). This co-solvent simulates the lipid environment and reduces the water activity, and may cause the peptide to take up its *in vivo* conformation. In the trifluoroethanol and water mixture, an α-helix which is marginally stable in pure water can be stabilized[5-7]. It has been shown that, in several peptide systems, TFE does not induce secondary structure indiscriminately, but regions of the polypeptide chain that are helical in the native state are stabilized preferentially[5-7]. Recently, the effect of TFE on

the structure of intact hen lysozyme in its native and denatured states has been studied. It was found that TFE stabilizes predominantly native-like secondary structure in a partially folded state of an intact protein[8]. This finding is of considerable significance and provides support for the idea that studies of peptide fragments of proteins in TFE can have direct relevance to understanding the structure and stability of protein folding intermediates.

Little is known as to why TFE stabilizes α-helices. The dielectric constant of TFE is 26.67, about one-third of the value, 78.54, for water at 25°C. Interactions between charged species would be expected to be stronger than in pure water. However, it was found that the lower dielectric constant of TFE was not important in the stabilization of α-helices in a peptide since the magnitude of the effect of charged groups on the peptide was not altered significantly by either TFE concentration or temperature[6]. Also compared to water, TFE is a much weaker base (pKa_1 ~ -8.2 for TFE vs ~ -1.8 for water)[9] and a stronger acid (pKa_2 = 12.4 vs 15.3)[9]. This means that TFE is stronger at donating protons for hydrogen bonds but weaker at accepting protons in hydrogen bonds. The effects of solvent acidity and basicity on peptide amides have been probed by NMR[9]. Hydrogen bonding from the peptide amide proton to a solvent acceptor results in decreased electronic shielding of the proton, causing a low-field shift for the proton. Hydrogen bonding from a solvent donor to the amide carbonyl also deshields the proton indirectly. Upon changing the solvent from water to TFE, an amide proton exposed to the solvent moves to higher field, because of the increased shielding due to reduced hydrogen bonding from the NH to the less basic TFE, while a carbonyl exposed to the solvent causes the NH to move to low field, because of decreased shielding due to increased hydrogen bonding of the more acidic TFE. When both NH and C=O are exposed, the effect of the basicity on the NH predominates. Overall, this could increase the ability for the

peptide to form intramolecular hydrogen bonds, thus stabilizing the α-helices[6]. Other possible stabilizing factors include lowering the chemical potential of water, binding of TFE to the peptide, and changes in the water structure.

In this thesis, the solvent used in all the studies is a mixture of methanol and water (60/40 v/v or 80/20 v/v). The physical properties of methanol are very similar to those of TFE. The dielectric constant of pure methanol is 32.63 at 25°C. This solvent may stabilize β-turns in small peptides[10-12]. Based on our studies on the peptide DEKS[11], ongoing studies on the FDEKA and YDEKA in both water and water/methanol solvents, evidence shows that methanol enhances the β-turn which exists in the pure water. Since the peptides we studied usually involve charged residues and higher concentration of methanol (80/20 or 60/40) were used in the studies, interactions between charge species seem to be important[11,12].

It has been suggested that important criteria for a useful structure stabilizer are[6]: (1) the transition from coil to folded structure should be cooperative as the stabilizer concentration is increased; (2) the importance of one type of interaction, such as the effects of charged groups, should not be predominantly amplified; (3) inherently more stable structures should be stabilized more than less stable structures; (4) the stabilized structure should be the same as that formed in the absence of the stabilizer. According to these criteria, the effectiveness of methanol as a stabilizer should be further tested, for example, by methanol concentration titration on the peptides which form parts of proteins and were found to adopt a higher percentage of β-turns in water[13,14].

The most difficult aspect of the study of peptides in solution results from the flexibility of peptides. Most proteins and peptides are flexible, with a dynamic equilibrium between various local energy minima. It is known that conformational flexibility is of importance for biological function and for folding

processes that proceed through numerous conformations spanning a relatively large phase space volume. For peptides, it is hardly possible to find a structure consistent with all measured NOEs/ROEs and the holonomic structural data (bond lengths and bond angles). It is likely that several dynamically interchanging conformations are involved. On the other hand, the finding of a structure which fulfills all constraints does not prove the absence of multiple conformations[16]. Usually, conformational processes with exchange lifetime τ_e in the range $10^{-4} < \tau_e < 10^{-1}$ s can be investigated by line shape studies. Faster intramolecular processes may be monitored through relaxation measurements. The effects on relaxation are dependent on the ratio τ_e/τ_c where τ_c is the overall molecular tumbling correlation time. For $\tau_e/\tau_c >> 1$, population-weighted conformationally averaged relaxation and cross-relaxation rates are observed. For $0.1 < \tau_e/\tau_c < 10$, intramolecular dynamics becomes directly relaxation-active, while for $\tau_e/\tau_c << 1$ conformational averaging of the molecular structural parameters occurs that depends on the details of the motional process. In addition, it is possible to sense, by rotating frame relaxation measurements, processes that modulate the chemical shift in the range $10^{-6} < \tau_e < 10^{-3}$ s[16]. Only in the case where exchange processes are slow compared to the molecular tumbling correlation time τ_c, i.e. $\tau_e/\tau_c >> 1$, and where population-averaged cross-relaxation rates are observed, can NOEs/ROEs have direct conformational implication.

In many cases, an assumption that one distinct conformation coexists with a random coil ensemble is made. Under this assumption, the distinct conformation (β-turns and α-helices) with a significant population and some characteristic NMR parameters can be detected by a set of criteria[3]. Among these criteria, medium range NOEs ($d_{\alpha\beta}(i,i+3)$ for α-helices, $d_{\alpha N}(i,i+2)$ for β-turns) are most important, since they indicate an overall folding within 3 or 4

residues. Coupling constants and temperature dependent coefficients are rather local indicators within the residue. Usually it takes the combination of unique NOEs, coupling constants and temperature dependent coefficients to confirm the existence of a secondary structure.

Quantitation of this population is not so straightforward[3]. The first problem is to define the boundary of the structure. It is not clear how large the range of torsion angles of the conformations must be in order to give rise to the medium range NOEs. Secondly, there is no way to tell whether all the characteristic evidence is cooperative. For example, a series of observed $d_{\alpha\beta}(i,i+3)$ NOEs indicate the presence of a α-helix along the peptide chain, but this does not mean that the α-helix is present simultaneously all the way along the peptide. A small coupling constant indicates that the ϕ angle is in the α-helix or β-turn region, but does not mean that the whole helix or turn is formed. Finally, certain information can be common to several conformations. For example, a type II β-turn has a strong $d_{\alpha N}(i, i+1)$ NOE, as does an extended conformation.

Attempts have been made to measure the population percentage of the secondary structure in multi-conformer peptides using NMR parameters, making the first-order approximations of uniform correlation times, idealized secondary structure geometry and isotropic motion (Brockbank, R. L., Ph.D thesis, U. of Calgary). The population of the secondary structures by this approach is believed to be greater than would be expected given a random distribution through the ϕ-ψ torsion space available in a Ramachandran map[3].

Recently, a search algorithm was presented by Ernst and his coworkers[16]. This algorithm allows the determination of multiple conformations of biomolecules in solution with exchange rate constants typically between 10^3 and 10^7 s^{-1} on the basis of experimental high-resolution data. Multiples of

structures are generated which are consistent with the ensemble of NMR cross-relaxation rates (NOESY, ROESY), homonuclear and heteronuclear J-coupling constants and longitudinal relaxation time $T_{1\rho}$ measurements. Another approach is to apply restrained molecular dynamics (rMD) for dynamic structures which has recently been proposed by Torda et al[17]. It uses a time-variable NOE restrained potential that depends on the past of the trajectory. These approaches probably reflect the trend in the conformational studies of small peptides in the future, which involves intensive NMR measurements (NOEs/ROEs, homonuclear and heteronuclear J-coupling constants, longitudinal relaxation time $T_{1\rho}$ measurements, and their changes with temperatures or field strength) and theoretical or empirical calculations (energy minimization and restrained molecular dynamics calculations).

For biologically important peptides, the most important question is whether the conformation derived from NMR or other methods is relevant to the conformation which is bound to a receptor. Unfortunately, determining relationships between structure and activity in small peptides is an elusive goal. Most of the biologically active peptides whose activities are well defined and measurable are highly flexible molecules. Despite the availability of numerous analogs and data on their respective activities, the number of possible conformations of these peptides complicates efforts to relate structural parameters and activities. Furthermore, the conformation observed in a particular environment suitable for physical measurements (most commonly bulk solvents) may well not be the receptor-bound conformation. The situation is complicated further by the likelihood that a conformational change on the part of the peptide molecule is a necessary aspect of its biological action[18,19].

Nonetheless, a great deal of effort has been invested in studies of preferred solution conformations of biologically active peptides. There is, no

doubt, a distribution of conformations in solution, and among the various conformers it is likely that the conformer that mediates biological activity will be found. But no physical method can yield a description of a single conformer among a population of conformers that interconvert rapidly on the time scale of the measurement. Moreover, there is no *a priori* strategy that would select the bioactive conformation from the population. Until a receptor can be purified and isolated and methods can be developed to examine the bound peptide molecule, only indirect approaches can be taken to determine the bioactive conformations of flexible peptides.

An approach that has proved useful in elucidating structure - activity relationships of bioactive peptides is the development of conformationally constrained analogs that retain the biological activity of the native molecule. The limited conformational space that is available to these constrained analogs must include conformers that interact productively with the receptor[18,19]. In fact, the conformation of the bound peptide can be studied directly by using the transferred NOE experiment under proper conditions[20,21]. In a mixture of bound peptide with 5-20-fold excess of free peptide, the peptide in the bound state may adopt a different conformation from those in the free state. By the exchange between the macromolecular binding site and the free state, the NOEs characterizing the conformation of the bound molecule can be transferred to more easily observable resonances, which correspond either to the spectrum of the free molecule modulated by the exchange with the macromolecular binding sites, or to the average of the spectra for the bound and free molecules, provided that the dissociation rate is fast enough. The problems with this method may come from the availability of the receptor and solubility of both the peptide and its receptor. An even more elegant way of studying the bound

conformation is to isotopically label the peptide and study the complex directly, using heteronuclear editing techniques[22].

For the collagen telopeptides which have been studied in this thesis, some preliminary transferred NOE experiments have been carried out in our lab, but a complete analysis could not be carried out since the collagen triple-helix fragment is not soluble enough to give reasonable sensitivity and the mixture is also not stable.

New, more potent bradykinin antagonists have been synthesized continuously by Dr. John Stewart. Several new antagonists are listed in Table 7.1 and their pharmacological data are presented in Table 7.2. It will be interesting to study their conformations in the solution, such as methanol/water, TFE/water or micelles, to test the hypothesis proposed in Chapter 5 and to provide more experimental evidence on activity-conformation relationship of bradykinin antagonists. As pointed out in Chapter 5, it is impossible to carry out NMR studies when these antagonists are bound to the receptor. These antagonists function at nanomolar concentration, far below the limit of NMR detection sensitivity.

References

1. Clore, G. M. & Gronenborn, A. M. (1991), *Prog. Nucl. Magn. Reson. Spectrosc.* **23**, 43-92.
2. Johnson, Jr., W. C. (1990), *Proteins* **7**, 205-214.
3. Williamson, M. P. & Waltho, J. P. (1992), *Chem. Soc. Rev.* **21**, 227-236.
4. Dyson, H. J. & Wright, P. E. (1991), *Annu. Rev. Biophys. Biophys. Chem.* **20**, 519-538.
5. Dyson, H. J., Merutka, G., Waltho, J. P., Lerner, R. A. & Wright, P. E. (1992), *J. Mol. Biol.* **226**, 795-817.
6. Nelson, J. W. & Kallenbach, N. R. (1986), *Proteins* **1**, 211-217.
7. Nelson, J. W. & Kallenbach, N. R. (1989), *Biochemisty* **28**, 5256-5261.
8. Buck, M., Radford, S. E. & Dobson, C. M. (1993), *Biochemistry* **32**, 669-678.
9. Llinás, M. & Klein, M. P. (1975), *J. Am. Chem. Soc.* **97**, 4731-4737.
10. Otter, A., Scott, P. G. & Kotovych, G. (1988), *Biochemistry* **27**, 3560-3567.
11. Otter, A., Scott, P. G., Liu, X. & Kotovych, G. (1990), *J. Biomol. Struct. Dyn.* **7**, 455-476.
12. Liu, X., Scott, P.G., Otter, A. & Kotovych, G. (1990), *J. Biomol. Struct. Dyn.* **8**, 063-080.
13. Dyson, H. J., Rance, M., Houghten, R. A., Lerner, R. A. & Wright, P. E. (1988), *J. Mol. Biol.* **201**, 161-200.
14. Dyson, H. J., Rance, M., Houghten, Wright, P. E. & R. A., Lerner, R. A.(1988), *J. Mol. Biol.* **201**, 201-218.
15. Braun, W., Wider, G., Lee, K. H. & Wüthrich, K. (1983), *J. Mol. Biol.* **169**, 921-948.
16. Brüschweiler, R., Blackledge, M. & Ernst, R. R. (1991), *J. Biomol. NMR* **1**, 3-11.

17. Torda, A. E., Scheek, R. M. & van Gunsteren, W. F. (1990), *J. Mol. Biol.* **214**, 223-235.

18. Rose, George D., Gierasch, Lila M. & Smith, John A. (1985), *Adv. Protein Chem.* **37**, 1-109.

19. Degrado, William (1988), *Adv. Protein Chem.***39**, 51-124.

20. Wüthrich, K. (1986), *NMR of Proteins and Nucleic Acids*, New York, Wiley.

21. Campbell, A. P. & Sykes, B. D. (1991), *J. Mol. Biol.* **222**, 405-421.

22. Muchmore, D. C., McIntosh, L. P., Russell, C. B., Andersonm D. E. & Dahlquist, F. W. (1989), *Methods in Enzymology* **177**, 44-73.

Table 7.1. A List of New Bradykinin Antagonists.

I	$DArg^0$-Arg^1-Pro^2-Hyp^3-Gly^4-Cpg^5-Ser^6-$DCpg^7$-Cpg^8-Arg^9
II	$DArg^0$-Arg^1-Pro^2-Hyp^3-Gly^4-Thi^5-Ser^6-$DIgl^7$-Oic^8-Arg^9
III	$DArg^0$-Arg^1-Pro^2-Hyp^3-Gly^4-Thi^5-Ser^6-$DCpg^7$-Igl^8-Arg^9
IV	Dhq-$DArg^0$-Arg^1-Pro^2-Hyp^3-Gly^4-Cpg^5-Ser^6-$DCpg^7$-Cpg^8-Arg^9

Abbreviations:	Cpg	α-cyclopentaneglycine
	Igl	α-(2-indane)-glycine
	Oic	octahydroindole-2-carboxylic acid
	Dhq	dehydroquinuclidine 3-carboxylic acid

Table 7.2. Pharmacological Data for Peptides I-IV listed in Table 7.1.

Peptides	pA_2(Rat uterus)	pA_2(Guinea pig ileum)
I	7.4	7.0
II	7.8	irreversible binding
III	7.4	7.2
IV	7.5	6.3

pA_2 = negative logarithm of molar concentration of antagonist to give 50% inhibition.